The Use of Levelled Assessment Tasks and their Impact on Teaching and Learning in Science Education

Andrew John Chandler-Grevatt

Doctor of Education

University of Sussex

May 2010

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:……

I dedicate this thesis to my parents, Kathy and John.

Acknowledgements

There are many people, each in their own way, who have helped and supported me during my thesis. I want to acknowledge and thank:

The teachers and students that took part in my research.

Badger Publishing, namely David Jamieson and the team for agreeing to this research and allowing me to include some published materials.

Professor John Gilbert for inspiring me to embark on this doctorate. The amazing people with whom I have worked throughout this doctorate, who have helped me grow personally and academically, these particularly include Sue Clayton, Jane Creaton, John Crossland, Michelle Lefevre and Sue Robertson.

The staff at the University of Sussex that have supported me through the course, these include Dr. Pat Drake, Dr. Mairead Dunne, Professor Michael Eraut and Dr. Elaine Sharland.

My supervisor Professor Jo Boaler, for her continuous encouragement and inspiration.

My readers and proof readers, James Williams, Joan Williams and Dr. Pat Drake.

My husband Geoff for his unwavering support and understanding every step of the way, as well as the technical support.

UNIVERSITY OF SUSSEX

Andrew John Chandler-Grevatt

DOCTOR OF EDUCATION

The Use of Levelled Assessment Tasks and their Impact on Teaching and Learning in Science Education

SUMMARY

The use of Levelled Assessment Tasks (LATs) in secondary science in England has been increasing over the past five years in response to attempts to encourage more Assessment for Learning (AfL) strategies in the science classroom. This empirical study investigates *how* LATs are used by teachers and the extent to which such tasks support teaching and learning. An online survey of 106 teachers was used. It showed that teachers did find that the LATs supported their teaching using AfL strategies, but revealed that a majority of teachers do not use the tasks as formatively as they could be used. From the online questionnaire, a descriptive framework for how the LATs support teaching and learning is proposed. Five case studies where teachers used a LAT were observed. The data collected included a post-lesson pupil questionnaire, an interview with a group of pupils and an interview of the teacher. From these cases, a theory seeking approach to educational case studies through fuzzy propositions (Bassey, 1999) was used to develop a model of the relationship between teacher values and pupil values to assessment tasks. The fuzzy generalisations proposed from the case studies were that: (1) Teacher attitudes to the LATs may influence pupil attitudes to the LATs, (2) Teachers with a 'big picture of levels' may be more likely to use LATs formatively and (3) Teachers who engage pupils with the notion of 'levelness' may be more likely to improve conceptual development of pupils. The notion of 'levelness' is explored. This evolves into three issues being explored: whether grades should be shared with pupils, the LATs relationship with summative and formative assessment practices and why such tasks have become popular with science teachers. The latter is considered in the context of the current target-driven culture of schools in England. Finally, the future of assessments like the LATs is discussed in relation to current policy and recommendations for their use and development are considered.

Contents Outline

Contents

List of Tables

List of Figures

Abbreviations

AfL	Assessment for Learning
APP	Assessing Pupil Progress
AST	Advanced Skills Teacher
CASE	Cognitive Acceleration in Science Education
CHSC	Cams Hill Science Consortium
DfES	Department for Education and Skills
DfEE	Department for Employment and Education
DCSF	Department for Children, Schools and Families
EdD	Doctorate in Education
FAT	Formative Assessment Task
GT	Grounded Theory
INSET	In Service Training
IWB	Interactive White Board
KCFAP	King's College Formative Assessment Programme
LA	Local Authority
LATs	Level-Assessed Tasks
NCLD	National Curriculum Level Descriptors
NSS	National Secondary Strategy
QCA	Qualifications and Curriculum Authority*
TES	Times Education Supplement
TGAT	Task Group on Assessment and Testing
ZPD	zone of proximal development

**as of 2009, the QCA has become the QCDA, the Qualifications and Curriculum Development Agency*

Chapter 1 Approaches to Assessment for Learning in Science Education.

Leading up to the new millennium a review of educational research was published which was to become arguably the most influential paper concerning classroom practice for many years. The findings of the review of assessment practices and their impact on learning had far reaching effects in the academy (Gardner, 2006 p. 6-7), on educational policy (e.g. DfES, 2005a) and on classroom practice (e.g. Black *et al.*, 2003).

'*Assessment and Classroom Learning*' by Black and Wiliam (1998a) brought to the forefront of the consciousness of academics, politicians and pedagogues, both the poor classroom practice that was in existence and the potential of alternative teaching strategies on improving learning in the classroom. Although the authors referred to these improved practices as 'formative assessment,' the policy makers encapsulated the principles for practitioners as 'Assessment for Learning' (AfL). In the decade since, the field of AfL research has grown and an abundance of teaching and learning strategies has developed.

Practising teachers, in a target driven culture, were encouraged to use AfL practices to improve teaching and learning and ultimately meet targets. Science teachers particularly, started to develop activities and approaches that could embrace some AfL principles.

My thesis is based on an empirical study of the use of one particular approach to improving AfL approaches in secondary science classrooms set in the context of policy and educational change.

1.1 A Professional Perspective

A significant theme of my thesis is that this project is the major part of a Professional Doctorate in Education (EdD). When I started the course, I was a full time Advanced Skills Teacher (AST) with a particular interest in AfL in secondary science and as a whole school initiative. I approached the National Secondary Strategies' (NSS) AfL drive from a pragmatic, practitioner, problem solving perspective. The tasks that I developed to support AfL strategies in science teaching and learning were a solution to a professional problem.

These circumstances have undoubtedly shaped my research design and so the narrative thread of this thesis follows my professional development and my progress as an apprentice researcher and academic. I have written this thesis to represent my perceived path of intellectual development with respect the practical problems I was facing as a science teacher. As such, the nature of my research has been theory-seeking and only as themes developed did I draw on an increasing amount of relevant research literature.

1.2 Assessment for Learning: solving a professional problem.

The substantive content of my thesis is based on a single type of classroom task, the so called 'Level Assessed Task' (LAT), which was designed to support both the teacher and the pupils' use of AfL in the secondary science classroom in England.

At the start of this research, I had already been a teacher for nine years and had written and published a number of books of LATs through Badger Publishing (e.g. Grevatt, 2005a, Grevatt, 2005b, Grevatt, 2005c). I was aware of, and had engaged with AfL as a concept in my practice, through policy (mostly via the National Secondary Strategy) and through some academic research (Black and Wiliam, 1998a).

In this section I explore the research, policies and practices that I was aware of in general and specifically in secondary science education. Then I focus on two particular aspects of AfL: the use of criteria based assessment and the approaches that teachers are using to understand the National Curriculum Level Descriptors (NCLDs).

1.3 Assessment for Learning: Practice, Policy and Research

Assessment for Learning (AfL) is a phrase that is now used in theory, educational policy and in secondary schools in the UK. For me as a school teacher, it was one of those 'buzz words' that cropped up within my workplace. It was not until I attended a NSS local authority training session in 2004 that I was introduced to 'Inside the Black Box' (Black and Wiliam, 1998b).

At that time, I was aware that cynical teachers saw AfL as another thing they have to do or be able to meet the expectations of performance management or Ofsted inspections. More informed teachers were aware of the sound-bites that had been distilled from the NSS support materials e.g. 'mind the gap', or had taken on some strategies such as traffic light cards or the no-hands rule (DfES, 2005a). More enthusiastic teachers may have read 'Inside the Black Box' (Black and Wiliam, 1998b) or 'Assessment for Learning: Putting it into practice' (Black et al., 2003) and then a few of them were passionate about the underlying principles of AfL and were slowly adapting their practice.

The 'Black Box' refers to the classroom and the way that current policy treats it, relying on inputs (demands or resources) and outputs (pupils learn, better test results etc.) and not the processes within the classroom: help and support for the teacher in the everyday classroom of achieving better learning (Black and Wiliam, 1998b p.1). Black and Wiliam (1998b) point out that formative assessment can improve standards, that there is

room for improvement in classroom practice and that there are evidence-based strategies for teachers to use: raising self-esteem through improved questioning practices; integrating self and peer assessment strategies into classroom practice; planning opportunities for pupils to express their understanding; establishing thoughtful dialogue between teachers and pupils and the formative use of summative assessments.

From my notes on my original copy of 'Inside the Black Box' I was able to relate to what they call 'a poverty in practice.' I was encouraged by how the features of formative assessment that were described echoed with my own classroom practice through writing and using Formative Assessment Tasks (FATs).

For my lessons, I had already been developing tasks that could be assessed using the NCLD in science (DfEE, 2004). This was unusual at the time because most teachers regarded levels as numbers that were generated from topic tests and the Key Stage National Curriculum Tests. I was fortunate enough to have been involved with a team of teachers who carried out action research to solve practical professional problems in science education, the Cams Hill Science Consortium (CHSC, n.d.). Together we had unpicked the NCLDs, which had informed my work on my FATs.

It was not until I had ventured into the part-time EdD at the University of Sussex that I read the original paper by Black and Wiliam. My first assignment was to critique an academic paper and I chose 'Assessment and Classroom Learning' (Black and Wiliam, 1998a). I critiqued the paper within the parameters of the question:

> *What is its relevance to classroom practice and could the changes suggested by this research be implemented in the current education system in England?*

I wrote the critique in 2004 in isolation with acknowledgement of, but not with direct reference to, subsequent research. In retrospect this decision was sound but having since read the subsequent research, many of my questions have been explored.

At that time, my main conclusions were that the lack of theoretical base was the largest hindrance to formative assessment becoming commonplace in classrooms. I felt that the examples given in the paper did not always seem to match my professional experience. I was also concerned that teachers would require a reason to change their practice. The review illustrated improvements in learning, but teachers were, as far as they were concerned, already improving their results according to the government's summatively based measures. Teachers would not want to risk a change in practice unless they felt confident that their summative assessment targets could still be met. I saw that formative assessment did provide an idealistic alternative and many teachers were dissatisfied with the over emphasis on grades and outcomes. I stated that without a convincing ground theory, that the authors were expecting a leap of faith from practitioners. I concluded that further research was required to help researchers and teachers understand the underlying theories and thus reasons for a change in practice and that there needs to be a government directed change of emphasis in assessment practices for teachers to have motivation to change. This was addressed by a teacher-friendly booklet (Black & Wiliam, 1998b). In addition, I proposed that if teachers were convinced to use formative assessment in the context proposed, there were several areas that would need to be addressed to ensure that the change was to be successful. It would require significant time, training and solid examples of good practice to develop and integrate these concepts. Not only would teachers need to develop skills in questioning and verbal feedback, support would be needed in developing tasks and approaches to deliver this in *all* subjects. The authors addressed this through working with schools and conducting educational research to support the integration of AfL strategies (summarised in Black et al, 2003). I thought that without a shift in expectation by the government, teachers would remain resistant to formative assessment practices. This is

interesting since Black and Wiliam's work was integrated into the government's National Secondary Strategy (DfES 2005a), a rare direct integration of educational research into policy. If the expectations did change to favour formative assessment approaches, teachers would slowly, with support, integrate the principles into their teaching. In a few years, a change in pedagogy may be observed. The authors have since celebrated the impact of their research in the classroom (Black & Wiliam, 2003). However, as my thesis reveals twelve years on, formative assessment is still only found in pockets of excellence within schools, it is by no means embedded in all teachers' practice.

It is now five years since I wrote this critique and the paper and the subsequent research noted above have a major influence on this thesis.

1.4 Initiation and development of the Level Assessed Tasks

The concept of 'levelled-assessment' through the setting of open-ended tasks first came to my attention during my second year of teaching. I remember teaching a Year 7 mixed ability class and having to quickly set some homework at the end of the lesson as I had forgotten it was their day for homework. Considering the lesson was about particle theory, I set the pupils a task to draw a poster of an ice cube melting and explain it scientifically. The homework that I collected a week later was remarkably varied. It seemed to me to illustrate the diverse understanding within the class and I was able to match most posters to a NCLD. Initially I found that I could classify the posters into those who showed understanding of particle theory and those who did not. This was the idea of an abstract model being used, equivalent to the science National Curriculum Level 5. Some pupils drew a simple picture of an ice cube melting, labelling 'solid', 'liquid' and 'melting' appropriately, best fitting Level 3 descriptors. Others included a brief description, using appropriate key words, achieving a Level 4. There were also a

couple of pupils that had included sufficient detail to be awarded Level 6 in that they were able to explain the relative movements and arrangement of the particles and relate this to the properties of the solid or liquid state.

Most pupils appeared to reach their 'predicted' level and this led me to realise that I could use this type of work as an indicator for assessing pupils' progress. It was not until a few years later, when I started to consider AfL as a concept, that I was able to adapt these open ended tasks to more goal-oriented activities.

Originally I wrote a Level Ladder from which I could mark the pupils' work more consistently; this was based on the NCLDs. The switch came when I tried to write the descriptors in 'pupil-speak' so that they had a scaffold to reach their target level. This improved pupil response to the tasks because they were more likely to comprehend what was required of them for the task. At this point the tasks were still used in a summative way.

It was soon after that I started to use the tasks in a way in which the pupils started to focus on improvement targets and started to give them opportunity to improve their work during lessons. The improvement became the focus of the lesson.

1.5 The Structure of the Level Assessed Tasks

The LATs that were subsequently produced were based on scientific concepts from each of the 36 QCA science topics (DCSF, n.d). Each task follows the same structure (see Figure 1.1). The tasks are composed of a title, a brief context, task instruction, keywords and the Level Ladder. I describe each component below.

The Title: The title for most tasks is a stimulus question for pupils to answer through attempting the task. However, sometimes the title is a catchy title that might engage pupils e.g. Interplanetary Postcards or The Dodgy Barbeque.

The context: Some tasks require more context setting than others, the ice-cube example only required a brief introduction. Some however, require a short paragraph to contextualize the situation or problem.

The instructions: The task instructions tell the pupil what they have to do and this is often related to the descriptors in the Level Ladder. In addition the task can differ between the two level ranges of the task. The task for levels 5-7, for example, would emphasise the use of the particle model.

The key words: Key words support the pupils who attempt the task and there is an expectation that pupils working at level 4 should use key words appropriately and describe what they mean. In later editions, tasks pitched at pupils working at higher levels, had fewer keywords with the expectation that these pupils should be able to volunteer their own.

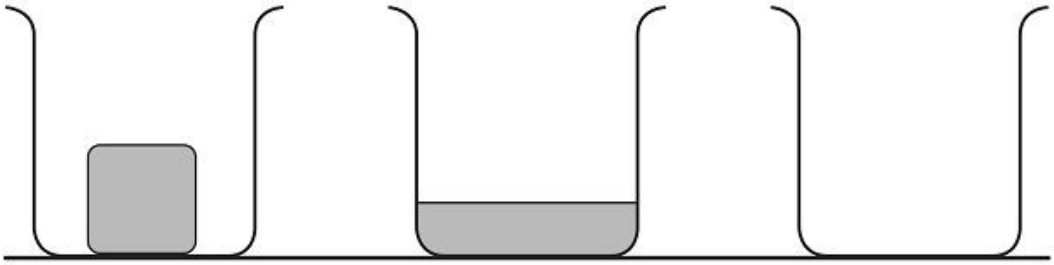
The Level Ladder: The Level Ladder is the component of the task that makes it unique. The Level Ladder is made up of descriptors that are written in language accessible to pupils. The NCLDs are quite generic and not accessible to most pupils. To overcome this, the level ladder contains statements that relate directly to the NCLDs with the addition of some generic statements. The statements are written using a combination of the general trends in the NCLDs and the sentences are structured using the stem words exemplified in Bloom's Taxonomy (Bloom, 1956, Anderson et al., 2001).

Figure 1.1 An example of a Badger Level Assessed Task.

7G Task Sheet (L5-7)

How does an ice cube melt?

Some students were watching an ice cube in a beaker as it slowly melted. They were wondering why it melts.



Task: Draw a poster that explains why an ice cube melts when left out of the freezer. Use the Big Idea of Particles to explain why the ice cube melts.

Key words:

boiling, compressible, conservation of mass, density, energy, evaporating, fixed, forces between particles, freezing, gas, liquid, melting, moving randomly, particles, solid, solidification, states of matter, temperature, vibrating

Level ladder:

What is your target level? Use the level ladder to help you reach it:

To get level	You might have:
5	• Used most of the key words accurately. • Drawn a simple particle diagram for each state. • Explained or shown that substances are made up of particles. • Described some differences between particle behaviour of each state.
6	• Used all the key words accurately. • Drawn particle arrangements clearly using diagrams. • Explained, in detail, the particle behaviour in each state. • Shown or described how mass is conserved during changes of state. • Explained evaporation using particle theory.
7	• Used a detailed scientific knowledge of particle theory. • Used energy and forces to explain the differences in behaviour of the particles in each state. • Explained the changes of state using particle theory. • Concept of energy and or forces should be incorporated into explanations.

From Grevatt (2005a) Task 7G

1.6 National Curriculum Levels in Science

These tasks were developed using the 1998 National Curriculum and are the focus of this study. The tasks have since been updated using the revised 2004 Level Descriptors.

The 1998 NCLDs were divided into four assessment areas or 'attainment targets'. The Department for Educational and Employment (DfEE) and Qualifications and Curriculum Authority (QCA) describe the attainment targets as:

> An attainment target sets out the 'knowledge, skills and understanding that pupils of different abilities and maturities are expected to have by the end of each key stage… attainment targets consist of eight level descriptions of increasing difficulty, plus a description for exceptional performance above level 8. Each level descriptor describes the types and range of performance that pupils working at that level should characteristically demonstrate. (DfEE, 2004 p. 74)

For science there were four attainment targets:

> Attainment Target 1: scientific enquiry
>
> Attainment Target 2: life processes and living things
>
> Attainment Target 3: materials and their properties
>
> Attainment Target 4: physical processes

To make some sense of these descriptors, teachers and educational professionals started to unpick the descriptors to get a 'generic feeling' for each level. This was achieved in several ways. I personally took the four descriptors from each attainment target for each level and drew out the key features of that level. This is what is often called 'levelness'.

For example Level 5 descriptors are:

> Attainment Target 1: scientific enquiry
>
> Not covered by these tasks
>
> Attainment Target 2: life processes and living things
>
> Pupils demonstrate an increasing knowledge and understanding of life processes and living things drawn from the key stage 2 or key stage 3 programme of study. They **describe the main functions** of organs of the human body [for example,

> the heart at key stage 2, stomach at key stage 3], and of the plant [for example, the stamen at key stage 2, root hairs at key stage 3]. They **explain how these functions** are essential to the organism. They **describe the main stages of** the life cycles of humans and flowering plants and point out similarities. They **recognise that there is a great variety of living things and understand the importance of classification**. They **explain that different organisms are found in different habitats** because of differences in environmental factors [for example, the availability of light or water].
>
> Attainment Target 3: materials and their properties
>
> Pupils demonstrate an increasing knowledge and understanding of materials and their properties drawn from the key stage 2 or key stage 3 programme of study. They **describe some metallic properties** [for example, good electrical conductivity] and **use these properties to distinguish** metals from other solids. They **identify a range of contexts in which changes** [for example, evaporation, condensation] take place. **They use knowledge about how a specific mixture** [for example, salt and water, sand and water] can be separated to suggest ways in which other similar mixtures might be separated.
>
> Attainment Target 4: physical processes
>
> Pupils demonstrate knowledge and understanding of physical processes drawn from the key stage 2 or key stage 3 programme of study. They **use ideas to explain how to make a range of changes** [for example, altering the current in a circuit, altering the pitch or loudness of a sound]. They **use some abstract ideas in descriptions of familiar phenomena** [for example, objects are seen when light from them enters the eye at key stage 2, forces are balanced when an object is stationary at key stage 3]. They **use simple models to explain** effects that are caused by the movement of the Earth [for example, the length of a day or year].

I have emboldened the phrases that could be described as the ‘generic’ phrases relevant to Level 5. The NSS also produced a ‘rough guide’ to levels (DfES, 2005b) which provides a ‘gut feeling’ for each level.

The QCA do not provide any explanation for the development of the attainment targets. Particularly for teachers, this appears to be a ‘black box’ itself since the QCA does not attempt to justify the theoretical background and wording for the descriptors. So teachers who are trying to make sense of the levels do this by a kind of reverse engineering. By looking for clues in the QCA descriptors, teachers have produced a range of ways of understanding the NCLDs for themselves and also for making these

accessible to students. The solutions that have been offered by science teachers include the ‘IDEAL’ approach, the ‘Levels Mountain’ and the Level Ladder. I discuss each of these below.

The educational theories that teachers seem to believe may underpin the science NCLDs are the developmental psychological theories proposed by Jean Piaget and the cognitive development theories of educational objectives developed by Benjamin Bloom.

The theory that childhood development goes through a series of sequential stages was proposed by Piaget (cited in Adey and Shayer, 1981) and has since been challenged (e.g. Donaldson, 1978) and developed in the light of new understanding and psychological research (e.g. Karmiloff-Smith, 1992). The ideas of Piaget have been used in curriculum design, particularly in mathematics and science education. Since Piaget’s focus was on cognitive development of concepts, both mathematics and science have concepts that teachers want to introduce and develop with their students (Adey and Shayer, 1994, Shayer and Adey, 2002).

Relevant to my research, is the work of Adey and Shayer that led to the classroom based activities and curriculum design (Cognitive Acceleration through Science Education). Before the introduction of the National Curriculum for England in 1988, Adey and Shayer had taken the key concepts from science and matched them to Piagetian phases of development (Adey and Shayer, 1981). Teachers have since interpreted this against the NCLDs, see Table 1.1.

This led to the scientific concepts met in science being matched to the Piagetian levels. However, Shayer and Adey highlighted that in the 1980s there was a mismatch between curriculum expectations and pupil readiness for concepts in science (Shayer and Adey, 1981 p.119). In their later work they acknowledge that the new National Curriculum

integrated some of the thinking skills and concepts, but by no means entirely (Shayer and Adey, 2002 p.1).

Table 1.1 Piagetian Levels of cognitive development

	Piagetian	My interpretation to levels
1	Pre-operational	
2A	Early concrete operational	
2B	Late concrete operational	Level 4
2B/3A	Transitional	Level 4/5? – needs support with abstract models
3A	Early formal operational	Level 5? – starts to use abstract models
3B	Late formal operation	Level 6 and beyond?

Adapted from Shayer and Adey (1981 p. 7)

In Table 1.1 the rough equivalence of the NCLDs and Piaget's phases are matched. I suspect that the NCLD for Science have been influenced by Piagetian thought and Adey and Shayer's research, although as I asserted earlier, for teachers this is not apparent. I have used the concrete-abstract divide as a framework when writing LATs and to help inform decisions about assessing pupils' work. Since teachers are not aware of the thought processes behind the assessment criteria, the drive of AfL has forced some to try to make sense of the NCLDs. An approach used by many science teachers has been to draw on Bloom's Taxonomy of learning objectives as a framework for making generic level-related assessments.

1.7 IDEAL approach

This is often seen in classrooms and on laboratory walls. It is a very simplified version of Bloom's Taxonomy being matched to the science NCLDs. This taxonomy is a framework for categorising educational objectives, first developed by Benjamin Bloom in 1956 (Anderson *et al.*, 2001). The word 'IDEAL' is an acronym of the key words associated with Bloom and these roughly match the NCLDs (Table 1.2). According to this, pupils who are able to *identify* a solid or a liquid are working at Level 3. Those working at Level 4 are able to *describe* what is happening, for example the ice-cube is melting. At Level 5 a pupil is able to *explain why* the ice-cube melts, and so on. Note that the Level 7 'link' refers to a pupil's ability to make links between two or more scientific ideas e.g. linking the particle theory to energy transfer to explain how an ice-cube melts.

Table 1.2 An example of IDEAL: Bloom and Levels.

Bloom Root Word	Level
Identify	3
Describe	4
Explain	5
Analyse	6
Link	7

This has advantages for classroom use in that it is very simple and catchy for both teachers and pupils. There are variations on this as well, some have 'IDEALS' with the 'S' relating to Bloom's word 'Synthesise' for level 8. However, I personally find it *too* simplistic. Although it may be accessible for pupils and teachers, it has stripped out the intricacies of Bloom's Taxonomy. There are two main opportunities that can be missed.

First, within a NCLD Bloom's stem words could be used with increasing demand for example, students working at Level 4 could have objectives using IDEAL. For example, the concept of burning magnesium could use Identify, Describe and Explain and Analyse within Level 4:

- Identify the type of reaction you see.
- Describe what you see when the metal burns.
- Explain what happens when the magnesium is put in the flame.

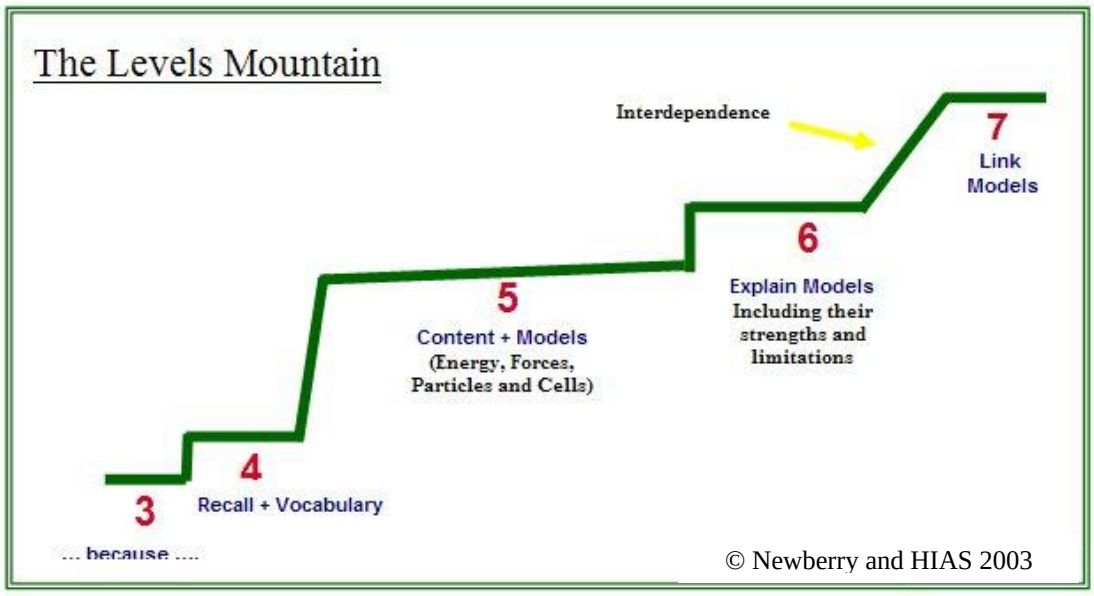
Secondly, when it comes to concrete and abstract concepts, Bloom's hierarchy of objectives can be of different demands depending on their context. In the case of the word 'describe', a pupil working at Level 4 could describe how an ice-cube melts, however, if the pupil could *describe* the arrangement of particles in a melting ice cube, this would be Level 6. The description of the ice-cube melting is a macroscopic phenomenon and therefore is a concrete concept, however, the concept of particles is an abstract concept, so although a low demand Bloom objective, when applied to an abstract concept, it arguably becomes more demanding.

1.8 The Levels Mountain

Through work on the development of teaching models and modelling (Gilbert and Boulter, 1998) some teachers engaged with the four 'key ideas' of Energy, Forces, Particles and Cells . This was developed as a framework within which any topic in science could be built from these key ideas. This was later adapted by the NSS to add 'interdependence' as a fifth 'key idea' launched in 2003. The next example of teachers making sense of NCLDs is the Levels Mountain. Making the descriptors pupil friendly involves conceptualising and, in the case of the Levels Mountain, visualising the nature of each level. Newberry, Gilbert and Hardcastle (2005) published this model, see Figure

1.2. This Levels Mountain was developed as a tool for both teachers and pupils to visualise their progress through the science Level Descriptors.

Figure 1. 2 The Levels Mountain



Through reference to five years' data from the National Tests, Newberry, Hardcastle and Gilbert (2005) contend that the journey of progression is not uniform: neither the journey through each level, nor the difficulty of transition between each level. The Levels Mountain was constructed through the use of quantitative data to estimate the relative height of the steps and the length of the horizontal slopes and by reducing the NCLDs for science into 'pupil speak' summary statements. They explain further:

> If the journey is imagined to be like climbing a staircase, you will notice that not all the steps are of equal height or depth. Level 5 is a key area since significant numbers of pupils fail to progress from level 4 at key stage 2 to level 5 at key stage 3… The first part of Level 5 requires pupils to be able to apply scientific ideas and explanations in unfamiliar situations whilst the second requires them to begin to engage with models of abstract concepts to explain phenomena. (p. 90)

Further evidence of the use and success of the Levels Mountain model has been published (Grevatt *et al.*, 2007) and the use of the model in conjunction with other

teaching tools. This attempt at visualising the NCLD, like IDEAL, can be seen on many school laboratory walls and often used as a way to communicate levels and progression in science with KS3 pupils. There are many variations of this model and sometimes it is merged with the IDEAL model.

The Levels Mountain has several advantages over the IDEAL model: it not only shows increasing challenge, the conceptual leaps needed between the levels, it also goes some way to explain why the 'journey time' through level 5 is longer than through level 4. It helps pupils (and teachers) to understand that levels are snap-shot assessments of work and that pupils can move up and down the levels and finally that the four key ideas can be used to explain any aspect of the science National Curriculum. The Levels Mountain does, however, rely on both pupils and teachers having a good understanding of the features of generic levels. This is partially why I developed the Level Ladders to communicate the requirements of each level between pupils and teachers.

1.9 Level Ladders

The concept of the Level Ladder was developed independently by other science teachers and indeed other subjects (e.g. Davies, 2002). This style of assessment is becoming increasingly relevant to teachers now that the Key Stage 3 National Tests have been abandoned (DCSF, 2008). Teachers are being left to use 'professional judgement' in assigning levels to their pupils' activities and work. In fact, it appears that the QCA and the NSS are endorsing this approach through the introduction of Assessing Pupil Progress (APP), which are essentially Level Ladders derived from NCLDs (QCDA, n.d.). This criterion-referenced assessment is becoming much more explicit in its use in the classroom, much more than the NCLDs have ever previously been used. This makes the findings of this thesis more pertinent.

1.10 Academic views on criterion-based assessment

I have presented three solutions that teachers have used to deal with using levels and this thesis will be focused on how teachers use the Level Ladder. Before addressing this, I will explore some of the academic literature with regard to NCLDs in light of these approaches.

Level descriptors are a form of criterion-referenced assessment. Sainsbury and Simar (1998) highlight the wide ranging interpretations and definitions of the concept of criterion-referenced assessment but assert:

> Criterion-referenced assessments, in an educational context, aim to give information about valued educational outcomes. They aim to tell us how well the pupils have learned what they have been taught. They are typically about cognitive outcomes, with understanding, knowledge and skill as central elements. This understanding, knowledge and skill is often described in fairly abstract and general terms. (p. 182).

This is as opposed to norm-referenced assessments which rank or compare one pupil with the others. This dichotomy sounds straight forward, but in practice, it becomes more problematic once teachers and examiners attempt to interpret the criteria when assessing student's work (Wiliam, 2001).

The academic literature contains views from people who have been directly involved with curriculum development (e.g. Cox, 1995) from subject specific subjects e.g geography (Davies, 2002) and the issues of criterion-referenced assessment (e.g. Sainsbury and Sizmur, 1998, Sadler, 2006, Sadler and Donnelly, 2006). It would seem that although teachers may have some faith in the integrity of the criteria upon which assessments are based, disappointingly this trust is flawed. Wiliam (2001) claims that the original recommendations were distorted by the individual subject groups before final publication of the NCLDs.

From an alternative perspective, much as teachers have been trying to do, Sainsbury and Simar (1998) have attempted some reverse engineering of the NCLDs under the assumption that:

> Our underlying assumption is that as a criterion-referenced assessment system, the National Curriculum was conceived with the intention that assessment results could be taken as providing accurate information about *pupils' attainment against the curriculum itself.* (p. 182).

They used a framework within which to analyse and critique the NCLDs for English, Mathematics and Science. They focussed on the relationship between what are called constructs, domains and performances, of which the criteria are composed. The authors state that the National Curriculum can be viewed as a set of educational constructs which are relatively complex, abstract and general in nature. The domains refer to the detail of what is to be assessed; essentially the function of the attainment targets (to become NCLDs in 1995). The performances are what are assessed by teachers in the classroom or by pupils' responses to tests. Sainsbury and Simar (1998) concluded that NCLD present considerable challenge to teachers:

> Teachers need to come to an understanding of the full nature of the educational constructs set out for them. In doing this, they will need to take note of the detail of the programmes of study and interpret this in the light of their professional subject knowledge. This interpretation, taken alongside the wording of the attainment targets, must give rise to an understanding of the nature of progression within the subject, as defined by the National Curriculum (p. 191).

Another matter that becomes pertinent at this point is the issues surrounding the 'abstract' concepts. Although the key abstract concepts have been highlighted by practitioners (e.g. Newberry & Gilbert 2007) and the National Secondary Strategy's 'Five Big Ideas' (DfES 2005b) and are used as a defining point of conceptual understanding at Level 5, it is apparent from the literature that these are far from understood. Extensive literature explores the teaching of particles. Black (unpublished)

explores this literature and exposes the lack of certainty on whether particles should be taught first when understanding changes of state. His analysis highlights that studies are focused on small qualitative samples and that a larger longitudinal survey would give more insight into average changes in progression in understanding.

The concept of energy was problematised by Robin Millar (2003) with a view to advise the National Strategies on an approach to teaching energy at Key Stage 3. This shows, as with particles, a lot of studies have attempted to elucidate how a concept of energy is developed. Again, there have been many small scale qualitative studies of individuals, including my own contribution (Grevatt, Gilbert and Newberry 2007). Here there is a philosophical debate between the energy transfer and energy transformation teaching models. Progression of the concept of force was explored by Bliss and Ogborn (1994) along with reasoning. Issues that arise for me at this point are that when writing Level Ladders it is with the view of scaffolding the 'next steps' within the demands of the national curriculum levels or to allow pupils to voluntarily come up with the key concepts. When writing them, my thoughts are with the former.

In summary, the publication of the Black and Wiliam paper presented an opportunity for policy makers and teachers to engage with AfL and its related pedagogies. In science education, teachers have attempted to embrace AfL strategies by engaging with the NCLDs. Science teachers have attempted to develop an understanding of NCLDs through a form of 'reverse engineering' to make sense of the criteria. The strategies that have come about from this are the 'IDEAL' approach, the 'Levels Mountain' and the 'Level Ladder.' It is clear from consulting some of the academic literature that the theoretical and logical integrity that teachers believe the descriptors have, does not exist. Furthermore, teachers' interpretation of the NCLDs against the work of their pupils is

varied. This thesis considers how the Level Ladder is used by teachers and pupils and the extent of its support for teaching and learning.

1.11 Research questions

This project has two stages through which the main questions were addressed. The use of LATs in science has had no published research associated with their use. So this research offers a starting point to explore how the tasks are used by teachers and pupils.

Before refining my research questions, it was important to get some insight into *how* LATs are being used by teachers and the *extent to which* teachers found the tasks supported teaching and learning.

My central research questions are:

How do teachers use the LATs?

To what extent do the LATs support teaching?

To what extent do the LATs support learning?

From these initial questions, further questions arose, some of which are addressed by this thesis:

To what extent do the LATs promote formative assessment practices?

How does the use of Level Ladders contribute to the debate about giving pupils their grades?

Why the LATs have become so popular with science teachers?

1.12 Originality of this Thesis

As far as I am aware, this is the first formal academic research into the use of Level Assessed Tasks in secondary science lessons in England. So in the first instance, my thesis provides a descriptive model of *how* this type of task is being used by teachers and learners and whether they are being used as I, the author of the LATs, intended. Subsequently, using a case study approach (Bassey, 1999), I have been able to offer a theory via a framework that is tentatively explanatory about the relationship between how the tasks are used and the resulting assessment, whether it is summative, formative or a mix. This forms a model of the relationship between how the tasks are used and the impact this has on teaching, learning and assessment culture.

From this I have been able to offer new contributions to the debate about the impact of giving grades to pupils and the effect on their motivation, an explanation for the popularity of the LATs with science teachers and an insight to the effects of performance demands on teachers and the unintended consequences of these demands.

Chapter 2 Methodology

2.1 An approach to conceptualizing methodology

Methodology and methods are distinct but intimately related components of educational research (e.g. Cohen *et al.*, 2000, Robson, 2002). Dunne, Pryor and Yates (2005) discuss the conceptualization of methodology as an 'elastic plane'. They explain that the shape of the methodology is produced by 'pulls' from six different dimensions of methodology (Figure 2.1). They offer this as a development of similar Venn diagrams used to represent this as they recognise that:

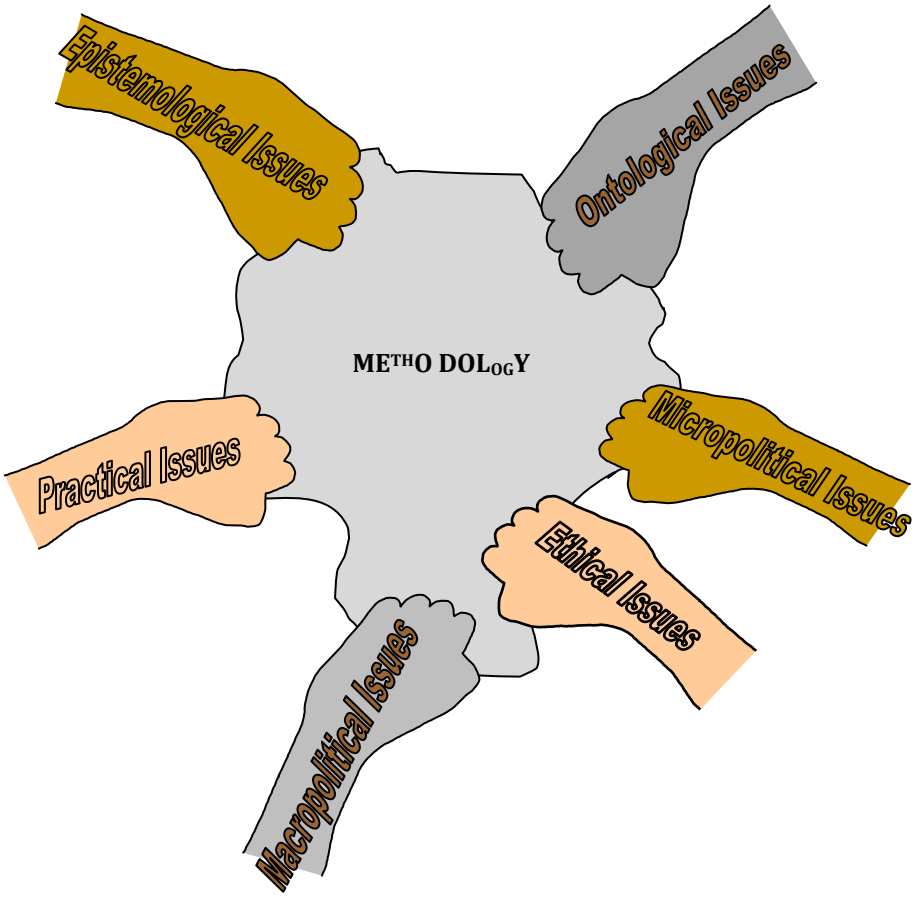
> social research does not take place in an idealized environment, but takes place in and investigates specific and changing contexts. In seeking for coherence, one may attempt to be consistent and to hold methodology still. However, contextual flux at all stages of the research process means that it is subject to pulling in different directions. The methodology of a piece of research is therefore liable to change its shape subtlety or sometimes quite markedly (Dunne *et al.*, 2005 p. 166).

This analogy is a useful framework on which to base a discussion of methodology. Therefore, I will discuss each component and then bring the issues together to define the methodological approach.

My research is concerned with finding out how teachers use a particular type of activity and the influence on their practice and their pupils' learning. To do this, I intend to use a mixed methods approach of a quantitative survey and five qualitative case studies. The data for the latter will be generated from lesson observations, short questionnaires and short interviews. The case studies themselves will be based on Michael Bassey's (1999) approach to educational case studies.

I will be collecting some facts about how the tasks are used for example when, how often and for how long. However, most of my data will be based on the users' perceptions, opinions and attitudes towards the tasks.

Figure 2.1 The methodological elastic plane (from Dunne *et al*., 2005)



2.2 Ontological Issues

When I was a natural scientist, ontological issues were not openly discussed nor were they thought to be relevant to be placed in a scientific report. Now studying the social world, it is important to understand the philosophical position from which one is conducting the research.

It is difficult to make claims about reality and although there are many attempts at categorising these into objective and subjective or positivist and interpretative or realist

and relativist (Blackburn, 2005), alternatively producing continua along which research or a researcher can position oneself (Cohen *et al.*, 2000 p.7), I have found that it is best to start from what one wants to achieve from the research.

For me, the purpose of educational research is to make a difference in the classroom either through improved understanding of pedagogy or pupils' learning experiences. In this research project I have chosen to collect evidence from direct observations and from teacher and pupils' views. From this, I am making an assumption that there is an objective 'real world' out there where the LATs are being used by teachers and pupils, who will offer me their subjective views about their experiences.

One of the reasons that I chose to use Bassey's (1999) educational case study as an approach was that he articulated the ontological position of a case study that resonated with my own position as a researcher:

> The public world is positivist; the private world is interpretative. That is a bold statement, but one that I suggest is a reasonable approximation to the truth...The problem for case study researchers is that when they unpick [these] simple positivist assumptions they find that concepts like 'school' and 'teacher' have a very wide range of meanings...The exportation of a particular case is essentially interpretative, in trying to elicit what different actors seem to be doing and think is happening, in trying to analyse and interpret the data collected and in trying to make a coherent report which is long enough to be meaningful and short enough to be readable. (Bassey, 1999 p. 44).

Ontology is complex and straddling paradigms does not sit easily with a natural scientist, however if I consider people's views as perceptions of 'reality', this approximates my experience of the world and equates with Bassey's bold statement.

2.3 Epistemological Issues

I have given the nature of knowledge a lot of thought while developing my research skills and engaging with educational research literature. As a natural scientist, the 'gold

standard' of double blind randomised trials were sought, however many of my small experiments did not meet this standard, but claims of knowledge could still be made. The gold standard was impossible to achieve in ecological settings because of the complex uncontrollable variables.

The same argument is made for social science methods, the complexity of social systems makes it difficult to conduct the scientific approach and even if it were possible, the ethical issues rightly override the quest for new knowledge (Hinchliffe and Woodward, 2004). A science teacher colleague once said to me that all educational research is worthless because unless it is a very large scale, randomised and statistically valid, it was not worth doing. As this was said at the start of my EdD, I have since given this view considerable thought.

The gold standard was not easily achieved in ecological settings, which is often an argument used in social research. RCTs are not regularly used in educational research. Robson (2002) summarises the key points of the significant and ongoing debate of the use of RCTs in social research. Appreciating the gold standard approach, he draws out four main obstacles to why RCTs are not the 'prime mechanism' of social science research: Systematic enquiry is a minor player in developing and changing society; many social researchers are apathetic to RCTs; RCTs are not feasible in a social context; social experiments, including RCTs, tend to yield equivocal results (pp. 116-117). For me, as Robson highlights, these are not reasons for avoiding the use of RCTs. However, further reading reveals that there is a history and loss of faith by some educational researchers due to previous attempts to make educational research purely experimental in the 1960s (e.g. Clegg, 2005 p. 417).

In more recent times, not all educational researchers are disillusioned with the approach, for example Torgerson and Torgerson (2001) propose a need for more RCTs in educational research, drawing on the evaluations of the National Strategy for numeracy and literacy. These were not done as RCTs and, they argue, are not robust enough to make firm conclusions due to factors such as the Hawthorne Effect and regression to the mean (pp.117-118). Clegg (2005) tackles this problem from a critical-realist perspective and although does not attack RCTs as a method *per se*, she unpicks the philosophical issues surrounding experimentation within a critical realist perspective. The arguments she explores are concerned with the sociopolitics to evidence-based approaches, how evidence is used and the epistemological dimension (pp. 417-418). For me, this appears relevant to my study because a RCT would not reveal the answers to my research questions. The focus is on gaining knowledge about *how* an intervention is used, rather than what the outcomes of that intervention are. However, if I was to evaluate the impact of the tasks on teaching and learning, an RCT would be suitable, providing I could have sufficient resources for a statistically sound sample.

Robson (2002) suggests that the way forward is to embrace the Pawson and Tilley (1997) approach to realistic evaluations, establishing 'what works, for whom, and in which contexts' (p.120), something which Clegg recognises with some caution (p.427). This is an alternative approach to evaluation, in my case the effectiveness of the tasks on improving teaching or learning.

As a classroom practitioner, I made decisions everyday about what worked and what did not during my lessons. Every hour that I taught could be seen as a mini case study on which I reflected and adapted my practice. Over the years this knowledge accumulated for me to become a proficient teacher and at that time was considered an expert in my field. None of this professional knowledge (Eraut, 1994) was obtained

through randomised double blind trials, but somehow I was making informed generalisations at the end of lessons, learning from the experiences and improving my practice. Within eight years I was regularly 'sharing good practice' and this professional knowledge was valued by my peers. At this point in time, I had already carried out several action research projects and supported colleagues in completing theirs. My professional knowledge was utilitarian, but it could not be said to be evidence based unless I had conducted some action research to reach that conclusion (McNiff, 1992).

The nature of my professional knowledge always was attempting to make generalisations from single cases, each case contributing to my body of professional knowledge (Eraut, 1994 p. 43). When supporting or advising other teachers, I would suggest strategies from my own experience that *may* work. The knowledge was always tentative.

This is second reason for using Bassey's framework. He integrates the concept of 'fuzzy generalisations' from case studies which provided me with a scaffold from which I could to propose generalisations from single cases. Bassey categorises generalisations from empirical study into three types: scientific, statistical or fuzzy. The first, he positions with the natural sciences where, he argues, cause and effect relationships are clear and laws can be determined. Bassey does not see this as part of social research; instead he offers the statistical generalisation from large samples as being more suited to complex social situations and works with probabilities rather than absolutes. Finally fuzzy propositions are predictions from empirical studies that state that something *may* happen, but without measuring its probability (p. 44-46). I will discuss fuzzy propositions in Chapters 3 and 5. Although I am not going to critically engage with the origins of fuzzy logic (Kosko, 1994), I will discuss the use of the concept in this context in the final chapter.

2.4 Macropolitical Issues

What Dunne *et al.* (2005) mean by the macropolitical issues is only illustrated through examples rather that explicitly defined. However, it appears to relate to my reasons for researching the subject. The whole purpose for me doing this research is to gain answers about my own work that could potentially inform and eventually improve classroom practice. This has issues regarding my position as an author, my motivations, the bias I may have towards wanting my own work to be successful. The issues here are common with any researcher who is researching their own practice (Robson, 2002 p. 534).

The 'pull' on this particular part of the methodology is that I need to be self-aware of my own assumptions about the use of the LATs. The research questions are not about how successful the LATs are or impact on grades *per se*. Perhaps another independent researcher would be better placed to do such research. Instead, the research focus is on how the tasks are used by teachers and pupils, comparing this with how I expect the tasks to be used and seeking evidence for how the tasks may support or hinder teaching and learning. This positions me to be able to make value judgments as an author and a researcher, which I hope will contribute to the research of criterion-based assessment and ultimately classroom practice.

2.5 Micropolitical Issues

The theme of my position as a researcher and the possible tensions of researching my own work are also evident in the micropolitical. These issues are mainly concerned with the relationships between the researcher and those who are being researched, namely the teachers and the pupils. Here it is worth noting that as the author of the tasks, there could be influence on the evidence I collect. The teachers I observed may have perceptions of me as a visitor to the school, as an observer in their classroom and as the

author of the tasks they are using. It is impossible to predict what these perceptions will be or the extent of these on the outcomes of my observations and interviews. The effect, if any, is more likely to be greater on the teachers who participate in my research than the pupils, who see me as nothing more than another teacher visiting their school and asking them some questions. However, these issues are important to be aware of when I am observing and interviewing and I discuss the actions that I take to address these in the methods section.

2.6 Practical Issues

The practical issues that shape this research are the timeframe in which the research can be carried out, the time that I could actually get to observe teachers using the tasks and the opportunities to visit teachers using the tasks. In my planning I had considered watching different teachers using the same task, but this proved impossible to arrange within the timeframe in which I wanted to conduct the case studies. Limitations included the number of teachers I could observe and had the time to transcribe collate and analyse the data as well as what I could realistically include in a thesis of forty-five thousand words. There is no doubt that this research has been shaped by opportunity, serendipity and the practical limitations of time and resources, being a lone researcher.

2.7 Ethical Issues

The University of Sussex, like all research institutions, has a set of ethical guidelines that are consulted and acted upon appropriately if ethical issues are raised. According to this, my research does not have any unusual ethical complications. The main ethical issues concern the effects on the teachers, pupils and schools who take part in the research and my role as a researcher when I am author of the LATs. I took appropriate action to ensure that all participants gave informed consent to taking part in my research

and ensured confidentiality and anonymity of institutions and the participants. There are several issues, which were manageable, but specific to different parts of my research methods. These are discussed within the methods.

It is the norm in educational research that takes place in specific schools, to change the name of the school and the participants to ensure anonymity (Robson, 2002) and thus protect them from any unforeseen repercussions of claims made by the research. Interestingly, Walford (2005) highlights a possible conflict with the generalisation of the evidence made by the researcher and the reader:

> The fact that none of the research schools is identified, implicitly gives the writer and reader the chance to broaden the findings of each study beyond the situations investigated. It gives a spurious generalizability of time and space to the results of specific studies. (sic) (p. 90)

Although this is particularly pertinent with the nature of case studies, I have chosen to work to the status quo. However, I will consider this further in Chapter 7.

With respect to the issues concerned with me researching the impact of my own published work, I have taken every opportunity to distance myself from the research when formulating the questions and when analysing the data. These strategies are described and discussed in later chapters (see sections 3.7.2, 3.7.6 and 3.7.9).

Having used the elastic plane analogy for discussing my methodology, some of these issues are expanded upon in the methods chapter.

Chapter 3 Methods

The methods that I chose to use were selected to gain evidence to address my central research questions:

How do teachers use the Levelled Assessment Tasks?

To what extent do the LATs support teaching?

To what extent do the LATs support learning?

3.1 Online questionnaire

Before refining my research questions, it was important to get some insight into how LATs were being used by teachers and the extent to which teachers found the tasks supported or improved learning. To do this, I constructed a questionnaire that I had originally used to assess the teachers use of the tasks before leading training sessions.

3.1.1 Design of the questionnaire.

Cohen, Manion and Morrison (2000) explore the ethics, approaches to design, structuring and delivery of questionnaires. Following their guidance on their three stage process of 'operationalizing a questionnaire' (sic) (p. 246), I designed the online questionnaire.

The first stage involves clarification of the general purpose of the questionnaire and translation into 'a specific concrete aim or set of aims'. The general purpose of my questionnaire was for me to find out how the LATs were being used by science teachers and their opinions of the tasks as a result of using them. This was done in order for me to define suitable research questions for the case study part of my research. The specific aims of the questionnaire were to find out:

Section 1: How the LATs are used

Section 2: Quantitative questions about the impact of the LATs

Section 3: Qualitative questions about the impact of the LATs

Stage two, is the identification and itemisation of subsidiary topics that relate to its central purpose and this was done by identifying the specific questions that would give details about each specific aim. Finally, Stage 3 involves the identification and itemisation of subsidiary topics relating to specific information. This was done mostly by tick boxes. The questions are shown below with the tick box options listed.

3.1.2 Online questionnaire questions

Section 1: How the LATs are used

Tick box for each question: Always, Often, Occasionally, Never

I set the LATs as homework tasks.

I get my students to do the LATs in a lesson.

When I mark a LAT, I assign an improvement target.

When I mark a LAT, I assign a level and sub-level.

I give my students opportunity to act on their improvement target during a lesson.

I give my students opportunity to act on their improvement target as homework.

I get my students to self-assess the LATs.

I get my students to peer assess each other's LATs.

Open text box: If you use the LATs in any way not mentioned above, state it below.

Section 2: Quantitative questions about the impact of the LATs

Tick box for each question: Strongly agree, Agree, Neutral, Disagree, Strongly Disagree

- The LATs have improved my understanding of science national curriculum levels.
- The LATs have improved my teaching using Assessment for Learning strategies.
- The LATs have improved the learning of my students.

- My students have an improved awareness of their current level in science since using the LATs.
- The LATs have hindered my teaching of science.
- My students have an improved awareness of how to improve in science, since using LATs.
- The LATs have hindered the learning of my students.
- The LATs have improved my assessment of my students.
- My students enjoy doing the LATs.
- I enjoy teaching using the LATs.

Section 3: Qualitative questions about the impact of the LATs

Open ended-responses:

- To what extent has using the LATs improved your teaching? Please give specific examples if possible.
- To what extent has using the LATs improved the learning of your students? Please give specific examples if possible.
- What problems have you had with using the LATs? Please give specific examples if possible.
- How do you think the LATs could be improved? Please give specific examples if possible.

Section 4: Optional responses

Name, sex, number of years teaching, local authority and email address.

The first two sections were closed questions, with an optional open ended box for free-text, so that teachers could identify anything I had not thought of. Section 3 was open-ended, allowing teachers to express their views in writing.

3.2 Collection of results

The pilot survey was trailed on paper at an INSET session that I delivered in London for ten science teachers, in November 2007. From this, one alteration was made and an electronic version was produced using the online survey designer – www.surveymonkey.com.

The use of Survey Monkey had several advantages, as it helped prevent the common pitfalls of questionnaire design (Cohen *et al.*, 2000) combined with my pilot questionnaire that identified issues with clarity of the questions.

The link to this survey was initially sent out to science teaching colleagues in East Sussex, West Sussex, Brighton and Hove, Surrey, Kent. Colleagues were asked to pass the link onto other science teachers that they knew. This was successful in yielding fifty responses. Then in April 2008, the survey link was posted on the *Times Educational Supplement* (TES) science forum (www.tes.co.uk). This initiated a further 56 responses.

3.3 Analysis of the online questionnaire

The 104 online questionnaire responses were imported into an *Excel* spreadsheet for analysis. The quantitative sections were converted to percentages for comparison and the qualitative responses for section 3 were coded.

The codes were decided once the data had been collected, this was so that categories of answer could be derived directly from the respondents' text. Some responses were given up to four codes. The codes were revised twice more, merging related categories, until the categories were distinct.

3.4 School visits

Once the online questionnaire had been analysed I was able to consider some key questions that I wanted to address through case studies. The questionnaire fulfilled its role at offering a reasonable sample size of response from a range of teachers who used the LATs. However, it did not give any insights into exactly how the LATs were used by teachers and pupils within the classroom.

3.4.1 Selection of and preparation for School Visits.

The school visits were chosen based on people that I knew would be willing to be observed and who used the LATs. There were no particular selection criteria, except that I had to be able to get to the school and back in one day, the lesson that the LAT was being done matched with a day that I was available and that the teacher had some

time after the lesson to be interviewed. As it turned out, none of the teachers I observed were friends or even colleagues of mine, which I felt was advantageous.

I contacted the Head of Science Departments that I knew and asked if they would recommend teachers to be observed. One teacher contacted me via email to ask if there was opportunity to be involved in any research that I was doing. At all times I ensured that teachers did not feel pressured into participating. This was done by contacting them by phone and email independently of their Head of Department, giving them the option to withdraw at anytime and working to their agenda when I could. One teacher I chose not to observe because I felt that they were reluctant and were being pressurised to participate.

Once I had been given a contact, I sent the teacher an email (Appendix A) explaining the purpose of the observation, detailing the purpose of the observation and explaining what the data might be used for. Once the date had been confirmed, I sent another email detailing the observation schedule and explaining what to expect from my visit (Appendix B).

3.5 Case studies

In order to gather some rich qualitative information about how LATs were used, I chose to use a case study approach. Case studies are a particularly diverse method for gathering and presenting knowledge, which can be employed by a range of methodological perspectives. Cohen et al. (2000) propose that the case study is ‘most naturally suited’ to what they define as the ‘interpretative and subjective’ paradigm of educational research approaches (p.181). When addressing the literature regarding the methodology and methods of case studies it is clear that there is no commonly agreed definition of a case study in educational research (e.g. Bassey, 1999, Taber, 2000).

Yin's (1993) book appears influential in most discussions, which offers a comprehensive analysis of case study approaches, advantages and pitfalls. The focus of most debates seems to fall into the following categories: a) the definition of a case study, b) the function of a case study and c) to what extent the evidence from case studies is generalisable.

3.5.1 Defining the case study

The definition of the case study is a well rehearsed and ongoing discussion, so here I will define my case studies in relation to the literature. The reasons for doing the case studies were twofold as part of this research: to observe, in detail, teachers using the LATs and to gather data from a variety of perspectives from within the classroom to gain insight into the use of LATs, teaching and learning. It appears that most discussions agree that a case study is bounded in space and time, it describes a case and often the evidence for the case is collected through mixed methods and validated through triangulation (Yin, 1993, Bassey, 1999, Cohen *et al.*, 2000).

3.5.2 The function of the case study

Case studies are about communicating the particular about a single situation. Cohen *et al.* (2000) list possible functions of the case study as possible advantages from a number of authors. They include that case studies are 'strong in reality', emphasising the subtlety and complexity, can be generalisable with other similar cases, can be immediately interpreted by researchers and practitioners and put into action and they catch unique features that may otherwise be lost in large scale surveys (p. 184).

Bassey (1999), drawing on other research, identifies three categories of educational case studies, which he recognises have overlap. There are other attempts at categorisation of case studies beyond educational research (e.g. Stenhouse, 1988). The three categories that Bassey proposes are theory-seeking and theory-testing case studies, story-telling

and picture drawing case studies, and evaluative case studies. Within these categories, my case study approach is theory seeking with some evaluative aspects. It is theory-seeking in that I am attempting to describe and explain how the LATs are used by teachers. The cases are evaluative in the context that research will be formative in developing the use of the LATs in the classroom and summative in that it will make some statements about how the LATs are used by teachers and pupils (Bassey, 1999 pp. 62-63).

3.5.3 Case studies and generalisability

The most contentious aspect of the use of case studies in educational research is the extent to which that they are generalisable, if at all (Robson, 2002). In Chapter 2, I discussed my thoughts on the generalisation of professional experiences into my professional knowledge and how Bassey's (1999) use of the framework of 'fuzzy generalisations' that supports an approach to making claims to knowledge from case studies.

Bassey (1999) defines a fuzzy generalisation as 'the kind of statement which makes no absolute claim to knowledge, but hedges its claim with uncertainties' (p.10). Even more tentative would be 'fuzzy propositions.' The important aspect of Bassey's approach is that the case study, in its published form, invites others to replicate the case study. He accentuates the tentativeness of the claims that can be made by a case. Of educational case studies he states:

> It reports something has happened in one place and it may also happen elsewhere. There is a possibility but no surety. There is an invitation to 'try it and see if the same happens for you'. (p. 52)

3.6 Almost Grounded Theory

When considering the methodological approach of this research, I explored using a 'Grounded Theory (GT) approach' (Glaser and Strauss, 1977, Taber, 2000, Charmaz, 2006, Hammersley, 2008). It is an attractive approach that appeared in the first instance to provide the solution to what I was trying to do: make sense of how teachers and pupils use the LATs in the classroom.

Glaser and Strauss (1977) developed GT to bridge the gap between theory and empirical research. It is considered as the discovery from data systematically from social research (p. 2). GT is an approach that is seated in the interpretative paradigm of social research. Although it is often emphasised that GT is a purely qualitative endeavour (e.g. Charmaz, 2006), Glaser and Straus (1977) do not see this as a clash, instead their focus is on the verification or generation of theory and claim that both quantitative and qualitative data are necessary, used together to generate theory (p. 18).

The attractive qualities of this approach were that it is evidence-based, that it was a strategy through which I could compare case studies systematically (Glaser and Strauss, 1977) and a method through which generalisations could be made through individual cases (Taber, 2000). However, there are two main concerns that I had regarding using this methodology. The first was the role of the researcher. Glaser and Strauss (1977) emphasise that the researcher should 'wipe their mind clean' and use the data to identify 'emergent theory.' I am not able to claim a blank mind when embarking on this research. As the author of the tasks, I bring my own preconceptions, expectations and opinions to the use of LATs. I believe that these are important in my role as a researcher, but in the case of GT may impede the process. The second concern was the key aspect of the process is generating 'saturated data.' I do not think the scale of my project could reach that point, where no more new codes arise from subsequent cases.

Although Glaser and Strauss do acknowledge that this point may never be met, with my five case studies of different LATs, I could not envisage reaching data saturation. If the research had focussed on one LAT or one teacher using a variety of LATs with one class, there may have been more chance of this.

Although, there will be some elements of the GT tradition throughout my project, I decided that Bassey's (1999) use of fuzzy generalisations through educational research was more appropriate for my research in this instance.

3.7 The structure of the case studies

The case studies that I carried out were based on a single lesson carried out by a teacher who was using the task. The boundaries that I set were that the task had to be based in Key Stage 3 (11-14 year olds). The topic of the task was not relevant. I had considered observing a number of lessons where the same task was used, however the features that I wanted to observe such as the lesson structure, the pedagogic approach, the interactions of the pupils with each other and the teacher would be the same whichever task was used. The LATs are a resource that the teacher uses and they pedagogy they apply is unlikely to be different when 'doing a Badger' style task.

I approached each case study systematically, with specific quantitative and qualitative tools for gathering evidence. Each case study was composed of data from the following instruments:

- A lesson observation (field notes).
- An end of lesson short questionnaire completed by all pupils.
- A group interview of 3-5 pupils (15-20 minutes).
- An interview with the teacher (10-30 minutes).

These instruments are described and discussed below.

3.7.1 Lesson observation

Being in the lesson to observe how the LATs were used was an important but probably the most problematic elements of the case study. I could have interviewed teachers about how they used the tasks, but I wanted to see firsthand what teachers were doing both consciously and tacitly (Eraut, 1994). I wanted to see the task used in the classroom context to get the 'feel' for how the pupils and teachers interacted with each other and the task itself.

The skill of observing lessons is something that I had developed in my role as an AST. I had observed lessons both formally and informally a variety of secondary school teachers in a variety of contexts. However, often these observations were done using the criteria based on Ofsted's standards or a University's Initial Teacher Training standards. Undoubtedly these skills came in useful when going into a lesson to observe. In contrast, my preconceptions of how I believe the tasks should be used and what constitutes a good lesson had the potential to hinder my observation and cause me to miss important aspects of the lesson.

I gave the lesson observations some considerable thought before entering the classroom. My main focus was to look for evidence for the tasks supporting or hindering the teaching and learning. The central issue I had was what exactly would constitute evidence for the LATs influencing teachers and pupils. I developed a lesson observation per-forma on which to record my observations (Appendix C).

It was important for me to keep this part of the research in perspective. The observation itself had practical importance and I could observe and note 'facts' that are arguably more resistant to interpretation. For example, the structure of the lesson, the timing of each section of the lesson, the way the LAT was used as a teaching or learning aid.

Then, during the lesson I looked for particular incidences when I thought that the LAT was having some impact on teaching or learning. These incidences included when the teacher or a pupil made direct reference to the LAT and questions were asked in relation to the LAT. I also kept my mind open for any other incidences that may occur that I had not predicted.

I could have written extensive notes during the lesson, using a double column presentation of the observational data, writing interactions word for word, including notes of the context and making explicit inferences relating to each line of text as Torrance and Pryor (1998) did for their studies of AfL in the primary classroom (e.g. p.109). Alternatively I could have videoed the lesson and made extensive notes from the play back. However, neither of these was 'fit for purpose' in the context of what I was trying to find out, in the time I had and the fact that the observations were 'one-off.'

Post observation, I quickly typed the lesson observation up into two formats. One a narrative of the lesson and the other extracts from the observation notes that provided evidence for the following statements:

Evidence that the LATs supported teaching
Evidence that the LATs supported learning
Evidence that the LATs hindered teaching
Evidence that the LATs hindered learning

I typed the observation summary narrative and evidence for the statements on the same day or next morning, so that the data were still fresh in my mind. The summary and the evidence for statements were both subject to my interpretation. I did not feel that this was a major issue when it is considered in the context of the range of evidence that I collected for each case study, but it does call for a reflexive approach to handling the data. As Cohen *et al* (2000) conclude their chapter on observation:

> Observation methods are powerful tools for gaining insight into situations. As with other data collection techniques, they are beset by issues of validity and reliability. Even low inference observation…is itself highly selective, just as perception is selective…In this respect it has been suggested that additional methods of gathering data might be employed, to provide corroboration and triangulation, in short, to ensure that reliable inferences are derived from reliable data (p. 315).

3.7.2 Specific ethical considerations for observations

Discussion of the ethical dilemmas of observation is focussed mainly around the covert approach to observing (e.g. Cohen et al., 2000, p. 315). My role as an observer did not need to be covert; the teacher had invited me in and the pupils were introduced to me as 'someone interested in their science lessons'. However there were other ethical issues I needed to consider when observing. These are different for the pupil and for the teacher.

As an observer I was taking a non-interventional approach. I did talk to pupils at appropriate times during the lesson, while they were working. I asked them questions about their work and the decisions they made about what they included and chose to present. I also asked if they had learnt anything new during the lesson and if so, what and how.

Spradley (1980) may call this type of participation passive or moderate in that to the pupils I had a role as an observer. Someone who was watching them work may ask questions, but not a teacher who would tell them what to do, answer their questions and tell them off if they misbehaved. However, there could potentially be situations where I would have to take an active role as a responsible adult. It would not be ethically sound to sit and, for example watch verbal bullying, physical assault, the teacher behaving unprofessionally or health and safety contraventions in a science practical without doing something active. That would then make me an active participant in the lesson and my perceived role would change.

I had decided that I would not reprimand pupils or even give them a ‘teacher look’ for misbehaviour; I could leave that to the teacher. If I witnessed bullying or a similar incident I could inform the teacher at the end of the lesson, if the incident had not been noticed. However, if there was a serious safety issue regarding the students or teachers well being, be it a dangerous practical or an individual endangering others I would have to intervene. This would have consequences on the data that I collected, but it would be ethical to do something in these rare situations.

3.7.3 Pupil questionnaire

In the final five minutes of the lesson, the pupils were asked by their teacher to fill in a short questionnaire that I had designed and provided (Appendix D). The questionnaire was composed of seven statements to respond to using one of five tick boxes labeled Strongly Agree, Agree, Not Sure, Disagree and Strongly Disagree:

1. I enjoyed doing this task.
2. I used the Level Ladder to help me to do the task.
3. I found this task difficult to do.
4. The Level Ladder showed me what to do to improve.
5. I learnt something new while doing this task.
6. The Level Ladder helped me instead of asking my teacher.
7. Another pupil helped me when I got stuck.

At the end of the questionnaire there were two boxes for pupils to respond in free text. These two questions were:

- What did you like best about doing this task?
- What do you like least about doing this task?

The purpose of this activity was to get some insight into all the pupils’ experience of and attitudes towards the lesson and the LAT. I was particularly interested in whether

pupils felt that they had learnt something new during the lesson and what they used to help them during the lesson, be it the teacher, the Level Ladder or their peers.

I chose a questionnaire that would take pupils no longer than five minutes to complete at the end of the lesson. I did not want it to be too intrusive, but felt it would be valuable to probe the thoughts of all the pupils in the lesson. This could relate to the more detailed pupil interviews and aid triangulation.

I asked teaching colleagues to assess it for accessibility and readability. The layout of the questionnaire was important in making the activity seem accessible for pupils, so I decided to use a landscape orientation, which spread out the text and boxes making the sheet look less daunting. I also wanted the pupils to take the questionnaire seriously, so I made it look 'professional' and asked them to put their name on it, although they were told that their responses would remain anonymous.

3.7.4 Issues concerning the questionnaire

One child with a physical disability in one case study, communicated with a Teaching Assistant, who completed the questionnaire with them.

3.7.5 Pupil Interviews

After the lesson, as soon as the lesson had finished, three to five pupils were selected by the teacher to be interviewed by me. I asked the teacher to select pupils based on a) they would talk b) a range of abilities and c) a balance of gender. I did consider selecting pupils myself during the lesson that exhibited interesting 'critical incidents' such as the LAT being used or not being used. This approach risked selecting pupils who would not be willing to talk and also may have influenced my interview from my preconceptions. I audio-recorded the interview on a digital voice recorder, this I could save to my computer and transcribe after the interview.

To ensure that the interview stayed focussed on my main research questions, I put together an interview schedule. This meant that the interviews were semi-structured but I was prepared to go 'off-script' if I needed to. Before the interview began, I read through a checklist with the pupils, so they could understand the purpose of the interview (Appendix E).

At the start of the interview I allocated each pupil with an alphabetical letter to recognise their work and their voice on the recording. They had their work from the lesson with them, which usually was in the form of a poster (but not always).

Onto their work, I asked pupils to attach Post-it notes on the part they thought was their 'best bit.' This was useful in relaxing the children before the main part of the interview. I then asked them in turn, the following questions from the interview schedule.

Interview Schedule
Ask pupils to discuss the following statements in relation to their work in front of them. • What do you think is the best bit about your work? • Did you learn anything while doing the task, if so what and how? • What did you find difficult when doing the task? • What would you do next time to improve your work?

At the end of the interview I asked the pupils if they had any questions for me before thanking them for taking part. A couple of the teachers were present for the interview, at the pupil's agreement. No doubt this could affect pupil responses, but because the questions were not focussed on the teacher, but the pupils' work, I thought it would not be a significant issue but noted it as part of the interview.

3.7.6 Ethical issues with pupil interviews

The two main issues that arose in the pupil interviews were whether I really had the pupils' informed consent and what benefit did the interview have for the pupils involved. With regard to informed consent, the pupils were told at the start of the lesson that they may be selected for interview, but they had a choice whether they took part or not. Towards the end of the lesson or at the end of the lesson the teacher selected the pupils. From what I observed pupils were given the choice whether to take part and I recall one pupil had something else to do, so was excused with no fuss. The pupils were then given an opportunity to withdraw from taking part at the start of the interview. No pupils that I interviewed took this option and I took time to make eye contact with each one. However, I do think it would take a particularly confident pupil to say no to a visitor at this point. Having said that, from my experience of working with pupils of this age, I think I would be able to pick up on any pupil that felt under duress and if I had detected that I would have found a way to ensure that they did not have to participate. As it was, this was not an issue.

The second potential issue was whether the pupils would actually benefit from the experience of being interviewed. I had not planned for the pupils to gain anything specific from the experience and was concerned if they were losing part of their break time or I had caused them to be withdrawn from another lesson. However, the very process of reflecting on their work, talking with their peers about the lesson and having the opportunity could well have been beneficial to the pupils. Also, two pupils involved in one of the issues continued to talk to their teacher about the lesson after the interview was completed, I felt that this was a useful process for pupils to have the opportunity to talk to their pupils on such a detailed level.

3.7.7 Pupils' work

The work that the pupils produced during the lesson and that they used during the interview was photographed using a digital camera. One photo was taken with the Post-it Notes on it, and another without. The photos were coded with the pupil code. These formed part of the case study. The teacher and pupils gave verbal consent for me to photocopy or photograph the work.

3.7.8 Teacher interviews

Originally I had planned to interview teachers for about thirty minutes, but the reality was that I had to conduct the interview in whatever time the teacher had available, the shortest being fifteen minutes. Again, I designed an interview schedule from which to structure the interview (Appendix F).

Teacher Interview Schedule

- Can you give examples of where the Badger task supported your teaching?
- Can you give examples of where the Badger task hindered your teaching?
- Can you give examples of where the Badger Task supported learning?
- Can you give examples of where the Badger Task hindered learning?
- How did you use the Task to plan the lesson?
- How did you adapt the task, why?

The interview was semi-structured because I was able to follow points of interest from the interviewee's answers when I felt it necessary.

3.7.9 Ethical issues with teacher interviews

Apart from assured confidentiality and anonymity that have been discussed in Chapter 2, the teacher interviews did not raise the same issues as the pupil interviews. Teachers were able to give informed consent and were also able to control the time in which I was able to interview them. I felt more confident that the participants knew what they

were doing and why they were doing it than I did the pupils. The teachers also benefitted from the experience through being able to reflect on their practice. In addition, some teachers appreciated being able to converse with me about LATs.

3.8 Constructing the case studies.

As I have embraced Bassey's (1999) approach to educational case studies, I constructed the case studies based on his guidance. Since I have got five case studies, I will have a two step approach, the first to decide on the outcomes of each case and then to synthesise those outcomes into a theory.

Bassey describes seven stages to the construction of the case study (p. 66):

1. *Identify the research issue, problem or hypothesis*
2. *Asking research questions and drawing up ethical guidelines*
3. *Collecting and storing data*
4. *Generating and testing analytical statements*
5. *Interpreting or explaining the analytical statements*
6. *Deciding on the outcome and writing the case report*
7. *Finishing and publishing.*

Stages 1 and 2 I have described comprehensively within Chapter 2. For the collection and storage of data, the collection methods have been described. However, the storage has been done both on hard copy in a file and most of the data is collated onto a single spreadsheet for each case study. Within this spreadsheet, all raw data and analytical data are stored, so that anyone can go back and look at the original data and the coding and analytical decisions I have made. Stages 4, 5 and 6 are discussed in my chapters 4, 5 and 6 respectively. The only difference being, that I use the outcomes through fuzzy propositions of my five case studies to generate a theory from all of their findings.

3.9 Coding the interviews

I used a GT influenced approach to qualitative data analysis to 'discover' emergent themes from the teacher and pupil interviews. However, the questions raised from the online survey did inform some of the concepts that I was searching for. Corbin and Strauss (2008) define coding as taking raw data and raising it to a conceptual level (p. 66). I have used the three levels of coding approach to 'generate theory'; open coding to find conceptual categories in the data, axial coding to find relationships between the categories and then selective coding to conceptualise and explain the relationships between the core categories (Robson, 2002, Corbin and Strauss, 2008).

The open coding stage involved immersing me in the interviews and the resulting transcripts and listing any code, whether conceptual or descriptive that came to mind. I called these the 'themes' and attempted to assign a theme to each line of the interview (Column J). This was done in an Excel spreadsheet (Figures 3.1 and 3.2).

Figure 3.1 Spreadsheet of codes from Teacher Interviews

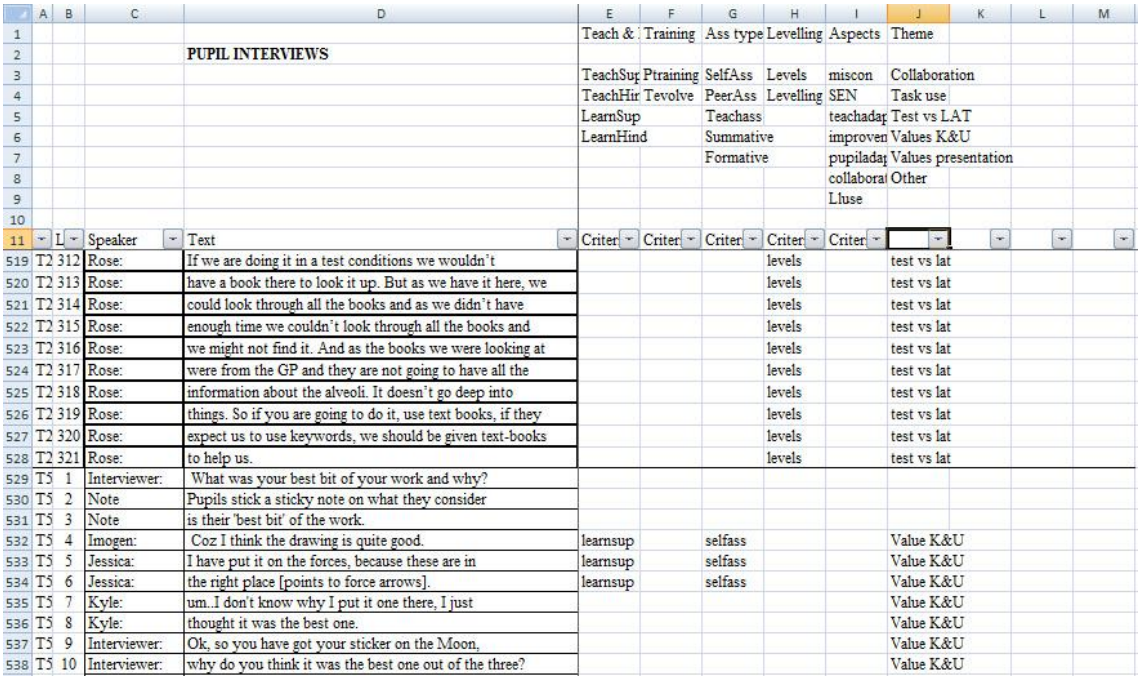
	A	B	C	D	E	F	G	H	I	J
1					Teach &	Training	Ass type	Levelling	Aspects	Themes
2				TEACHER INTERVIEWS						Benefits
3					TeachSup	Ptraining	SelfAss	Levels	miscon	Collaboration
4					TeachHir	Tevolve	PeerAss	Levelling	SEN	Copying
5					LearnSup		Teachass		teachadap	Classroom culture
6					LearnHind		Summative		Lluse	Getting confidence
7							Formative		Collabora	Learning from task
8										Time
9										Target setting
10										Values K&U
11										Values presentaion
12										Other
13		Lin	Spea	Text	Criter	Criter	Criter	Criter	Criter	
32	T1	19	T1	can't say it for every one. But because it is now	LearnSup					Classroom culture
33	T1	20	T1	routine to fairly regularly look at what level you	LearnSup					Classroom culture
34	T1	21	T1	are and improve it, then that ethos gets spread	LearnSup					Classroom culture
35	T1	22	T1	throughout the classroom, that becomes the kinda	LearnSup					Classroom culture
36	T1	23	T1	the of routine and we try to do in other lessons, we	LearnSup					Classroom culture
37	T1	24	T1	try to do it in every lesson. But I think that has	LearnSup					Classroom culture
38	T1	25	T1	stemmed from sort of that framework of using the	LearnSup					Classroom culture
39	T1	26	T1	tasks. And that's where I kind of got the skills and	LearnSup					Classroom culture
40	T1	27	T1	they have got used to that.	LearnSup					Classroom culture
41	T1	28	I	Can you point out any examples from today.						
42	T1	29	T1	Well even just things like pointing out to the	LearnSup					Classroom culture
43	T1	30	T1	kids erm they know I will put up a model now, so	LearnSup					Classroom culture
44	T1	31	T1	they will just kind of jot stuff down. But there	LearnSup					Classroom culture
45	T1	32	T1	were quite a few of them putting particle diagrams	LearnSup					Classroom culture
46	T1	33	T1	for conduction with the particles far apart. They	LearnSup					Classroom culture
47	T1	34	T1	know they are supposed to be together in a solid.	LearnSup					Classroom culture
48	T1	35	T1	But because they have copied…but actually the	LearnSup					Classroom culture
49	T1	36	T1	fact that they did it wrong meant it was easier to	LearnSup					Classroom culture

The selective coding came about through the open coding. However, the first category (Column E on the spreadsheet) was taken from the research questions. The criteria for evidence that using the LAT either supports of hinders teaching or learning (*TeachSup, TeachHind, LearnSup, LearnHind*). This was what I was looking for in my lesson observations; it could take any form from the teacher or pupil making reference to the Level Ladder to the teacher or a pupil expressing a like or dislike of using the task.

By identifying the emergent themes, I was able to start categorising the rest of the themes. This led to five categories that can be seen listed in Figures 3.1 and 3.2 (Columns E-I). The category of training was referred to formal or informal training or development of using the LAT; *Tevolve* referred to how the teacher's use of the LAT had evolved and *Ptraining* for any reference to how pupils learnt how to use the LAT itself.

The type of assessment that was taking place while using the task, be it summative or formative and/or teacher, peer, or self assessment (according to Black and Wiliam's, 1998 definitions) was recorded as the third category. The fourth category was whether I felt a particular statement or observation provided evidence for teachers or pupils having a notion of '*levels*' or '*leveling*'. '*Leveling*' I later changed to '*levelness.*' Finally 'Aspects' was an opportunity to code any concepts that I considered may be relevant during the coding. This section did go through several revisions to merge similar codes.

Figure 3.2 Spreadsheet of codes from Pupil Interviews



						Teach &	Training	Ass type	Levelling	Aspects	Theme				
				PUPIL INTERVIEWS											
						TeachSup	Ptraining	SelfAss	Levels	miscon	Collaboration				
						TeachHin	Tevolve	PeerAss	Levelling	SEN	Task use				
						LearnSup		Teachass		teachadap	Test vs LAT				
						LearnHind		Summative		improven	Values K&U				
								Formative		pupiladap	Values presentation				
										collabora	Other				
										Lluse					
	L		Speaker	Text		Criter	Criter	Criter	Criter	Criter					
T2	312		Rose:	If we are doing it in a test conditions we wouldn't					levels		test vs lat				
T2	313		Rose:	have a book there to look it up. But as we have it here, we					levels		test vs lat				
T2	314		Rose:	could look through all the books and as we didn't have					levels		test vs lat				
T2	315		Rose:	enough time we couldn't look through all the books and					levels		test vs lat				
T2	316		Rose:	we might not find it. And as the books we were looking at					levels		test vs lat				
T2	317		Rose:	were from the GP and they are not going to have all the					levels		test vs lat				
T2	318		Rose:	information about the alveoli. It doesn't go deep into					levels		test vs lat				
T2	319		Rose:	things. So if you are going to do it, use text books, if they					levels		test vs lat				
T2	320		Rose:	expect us to use keywords, we should be given text-books					levels		test vs lat				
T2	321		Rose:	to help us.					levels		test vs lat				
T5	1		Interviewer:	What was your best bit of your work and why?											
T5	2		Note	Pupils stick a sticky note on what they consider											
T5	3		Note	is their 'best bit' of the work.											
T5	4		Imogen:	Coz I think the drawing is quite good.		learnsup		selfass			Value K&U				
T5	5		Jessica:	I have put it on the forces, because these are in		learnsup		selfass			Value K&U				
T5	6		Jessica:	the right place [points to force arrows].		learnsup		selfass			Value K&U				
T5	7		Kyle:	um..I don't know why I put it one there, I just							Value K&U				
T5	8		Kyle:	thought it was the best one.							Value K&U				
T5	9		Interviewer:	Ok, so you have got your sticker on the Moon,							Value K&U				
T5	10		Interviewer:	why do you think it was the best one out of the three?							Value K&U				

The themes from the last column of the spreadsheet are displayed in Tables 4.9 and 4.10. These compare the occurrence of these themes in different case studies for teachers and pupils. It gives a flavour of the nature of each case, where cases share characteristics and where they differ. The codes were not cross checked by anyone else due to me being the lone researcher.

3.10 Discussion

The evidence I have planned to collect is through an online questionnaire to generate some quantitative and qualitative data about how LATs are used and teachers' perceptions of how the tasks may support teaching and learning. This sample of 106 teachers is then used to make some claims about the general use of LATs and to identify questions to research through five case studies. The five case studies are of single lessons where a teacher uses a LAT. These cases first generate a series of fuzzy propositions, which are then used together to generate a theory of LAT use and the

extent to which they support teaching and learning. The latter stage particularly has resonance with the GT approach, as discussed earlier.

Bassey (1999) defines the characteristics of an educational case study (p. 58):

- conducted within a localized boundary of space and time;
- into interesting aspects of an educational activity, or programme, or institution, or system;
- mainly in its natural context and within an ethic of respect for persons;
- in order to inform the judgments and decisions of practitioners or policy-makers;
- or of theoreticians who are working to these ends;
- in such a way that sufficient data are collected for the researcher to be able

(a) to explore significant features of the case,
(b) to create plausible interpretations of what is found,
(c) to test for the trustworthiness of these interpretations,
(d) to construct a worthwhile argument or story,
(e) to relate the argument or story to any relevant research in the literature,
(f) to convey convincingly to an audience this argument or story,
(g) to provide an audit trail by which other researchers may validate or challenge the findings, or construct alternative arguments.

These characteristics do resonate with most forms of qualitative research for example GT, naturalistic and ethnographic approaches (Cohen et al., 2000). Relating this back to the Dunne et al. (2005) elastic plane analogy (see Figure 2.1) it is mostly the practical, epistemological and ontological 'pulls' that shape the methodology of the research. From this it is difficult to see what makes Bassey's educational case studies different from any other type of qualitative research. The issues in defining case studies were explored in section 3.5. Bassey's second bullet point is probably the most defining of an educational case study in that it is specific to this approach.

My case studies do satisfy the first five bullets. Each study takes place in one classroom in one lesson, the interesting aspect is the use of LATs, the research is to be conducted ethically and the results it is anticipated will help inform judgements and decisions of practitioners at least and maybe contribute to the body of knowledge used by policy

makers and theoreticians. At this point in the thesis, the most pertinent point is (c) to test the trustworthiness of these interpretations.

Bassey does not accept that the terms reliability and validity as useful when dealing with case studies (p. 75). Instead he draws on the work of Lincoln and Guba (1985) who propose a concept of trustworthiness from a naturalistic perspective. Bassey synthesises this with his stages of a research project to produce a series of questions that are a summary from which discussions of trustworthiness of a particular case study can be built.

Stage 3: collection of raw data:
1. Has there been prolonged engagement with data sources?
2. Has there been persistent observation of emerging issues?
3. Have raw data been adequately checked with their sources?

Stage 4: analysis of raw data:
4. Has there been sufficient triangulation of raw data leading to analytical statements?

Stage 5: interpretation and analytical statements:
5. Has the working hypothesis…been systematically tested against the analytical statements?
6. Has a critical friend thoroughly tried to engage with the findings?

Stage 6: reporting of the research:
7. Is the account of the research sufficiently detailed to give the reader confidence in the findings?
8. Does the case record provide an adequate audit trail?

Adapted from Bassey (1999) p. 75

All these points have been or will be dealt with in the next chapters. The main issue that is important to discuss at this stage is the methods of triangulation from point 4. Bassey commits three sentences to the concept of triangulation (p. 76), stating that it a means of strengthening confidence in a statement. Next, I consider the concept of triangulation.

Triangulation can be defined as the use of two or more methods of data collection that in social sciences attempt to explain human behaviour from more than one standpoint (Cohen et al, 2000 pp.112-113). Gorard and Taylor (2004) discuss triangulation in

relation to ‘combining methods’ that in research, like mine, uses both qualitative and quantitative data. They propose a definition, in which they do not consider the notion of different research paradigms as a barrier, that involves a minimum of two vantage points or datasets telling us something about a third phenomenon.

A critique of the common analogy of triangulation from land surveying is given by Gorard and Taylor (2004), where they assert that it assumes that there is a genuinely stable phenomena that is being observed (this would be at odds, they argue, with both positivism and relativism). Second, that triangulation cannot be used as a form of mutual confirmation or validation of the two observations, because we are not able to tell if any of the points is an error. So, thirdly, they assert that triangulation is about ‘complementarity’ (sic) and nothing to do with validity (the latter supports Bassey’s claim). They conclude that the methods should be complementary, producing different aspects of reality under investigation and then put together (p. 46). The relevance of this to my study is about the claims I make, whether they are from the online questionnaire or from an individual case study. I will be attempting to produce fuzzy generalisations (tentative) or fuzzy propositions (more tentative) from individual case studies and the five case studies together.

I have planned to use multiple perspectives (teacher, pupil and researcher) and multiple methods, both quantitative and qualitative (questionnaires and interviews) from which I can triangulate. When triangulating claims, I envisage that I can determine the strength of the claim and make assertions of whether it is a statistical generalisation, a fuzzy generalisation or a fuzzy proposition.

Chapter 4 Data and Initial Findings

4.1 Initial research questions

Before refining my research questions, it was important to get some insight into *how* LATs were being used by teachers and the *extent to which* teachers found the tasks supported teaching and learning.

4.1.1 Online questionnaire findings

The purpose of the first section of the online questionnaire was to establish an insight to how classroom teachers typically use the LATs. The three areas of interest were:

a) How the tasks are typically delivered (during a lesson or as homework);
b) How the tasks are assessed (by the teacher assessment, peer-assessment or self-assessment and whether the teacher assesses using levels and improvement targets);
c) If the teacher provides opportunity for pupils to act on improvement targets.

4.1.2 Delivery of the tasks.

The data for this are presented in Table 4.1. The typical delivery of the tasks appears to take place in the lesson, with only five respondents never using the tasks in class. However, 32% of teachers never use the tasks as homework activities, 50% occasionally do and only two state that they *always* use the tasks in this way. The relevant additional comment made by respondents was:

> *We do them in the lesson and they are sometimes allowed to complete them for homework.*

The questionnaire did not allow teachers to state categorically if they allow pupils to complete the tasks for homework after setting the task during lesson time, but as this teacher states it is another approach. I had failed to give teachers the option to state if they used the tasks in test conditions, which since I have found that some teachers do attempt. This was a missed opportunity.

Table 4.1 How the Level Assessment Tasks are used.

	Always	Often	Occasionally	Never	Total
I set the LATs as homework tasks.	2	16	**52**	34	104
I get my students to do the LATs in a lesson.	32	35	32	5	104
When I mark a LAT, I assign an improvement target.	**63**	27	12	2	104
When I mark a LAT, I assign a level and sub-level.	**67**	15	14	8	104
I give my students opportunity to act on their improvement target during a lesson.	14	36	**45**	9	104
I give my students opportunity to act on their improvement target as homework.	3	31	**45**	25	104
I get my students to self-assess the LATs.	15	**40**	39	10	104
I get my students to peer assess each other's LATs.	18	**45**	32	9	104

However, the additional comments do give some idea of the ways in which tasks are used, particularly as an alternative to an end of topic test:

> *Use as an assessment at end of topic to supplement SAT-type test result.*
>
> *As an alternative to an end of topic test.*

Some teachers described specific approaches such as the use of ICT and the use of collaborative work, the latter not providing any detail of how this is carried out:

> *I tend to use ICT for them to produce a poster (usually via PowerPoint) of their work. This sometimes helps students with poor literacy skills. It also helps them to engage better with the task.*
>
> *Collaborative work.*

4.1.3 Assessment of the Tasks

With regard to the assessment of the tasks, it appears that 64% of teachers assign a level and sub-level to pupil's work. The questionnaire did not allow for teachers to specify if they set a level alone, but one teacher stated:

> *Give level but not a sub level. Only second year of using in dept.*

I am interested whether teachers are focussed on the assigning of a level to the work or engaging in the process of 'levelling' with their pupils. The following statement suggests that the teacher is focussed on the act of assigning levels – using the concept of 'levelness':

> *Moderation in dept meeting to ascertain teachers grasp of levelness and discussion of what a level 5 is and what good target comments look like.*

Another teacher appears to not use 'levels' themselves but uses the criteria alone to show pupils of how to get 'their target level':

> *I have set them as classroom tasks without leveling the outcomes - but using them so students can see what depth of knowledge corresponds to their target level.*

I am not sure how this works from the comment itself, but it appears the teacher is using the Level Ladder without the numbers. Although the predominant approach to marking the tasks is the teacher themselves, a vast majority of teachers do provide the opportunity for pupils to peer and self assesses. However, it is not clear from the data if this is summative or formative assessment.

4.1.4 Provision for improvement

In my view, providing pupils with the opportunity to improve on their first attempt is essential to providing formative assessment. However, this is the most interesting statistic because only a third of respondents often give their pupils an opportunity to

improve and a quarter of the teachers *never* provide that opportunity. This suggests that in the main that the tasks are often used summatively and are not used at their full potential as tasks that can support AfL.

A teacher states that:

> *I allow the pupils to select their own improvement target so that they recognise how to improve their own work.*

Although improvement targets appear to be given by a lot of teachers, pupils are not often given the opportunity to actually improve. This could suggest that although teachers are going through the motions of AfL, the actual potential of the process is not being reached and therefore reducing learning opportunities.

4.1.5 Summary of how the LATs are used.

Within this sample of 104 teachers who use the tasks, the following general statements can be made:

- Almost all teachers use the LATs during the lesson a few sometimes set them for homework.
- A majority of teachers provide the pupils with a level for the task.
- A majority of teachers provide improvement targets regularly.
- Over half of teachers regularly provide opportunity for self or peer assessment of the tasks.
- A minority of teachers regularly give pupils the opportunity to improve their work.

This summary could be represented as a 'pen portrait'[1] of how a typical teacher uses the tasks. This is where my views as the author of the tasks are potentially useful. This pen

[1] A term that I have borrowed from http://www.chrisg.com/how-to-create-pen-portraits-and-understand-your-target-audience/

portrait shows me that generally teachers are not using this tasks in the way in which they were intended. In the introduction to the *Badger Science Key Stage 3 Levelled Assessment Tasks* I describe how I expect the tasks to be used:

> These tasks are ideal to use either mid-way or towards the end of a topic. As the tasks have evolved and been trialled, many approaches have been tried. These are outlined below. Whatever approach you decide to use, make sure that the tasks are formative. It is important that these are not used as replacement summative tests. They are designed to encourage learners to demonstrate what they understand and to have the opportunity to improve. This is the foundation of formative assessment strategies: Where am I now? What am I aiming for? How do I get there? To aid this, learners must be aware of the level (and sub-level) for the end of the year.
>
> The tasks are designed to give learners the opportunity to show their full potential in science. To ensure this, I allow the class to use their notes from exercise books, text books and other secondary sources to help them with the task. I also encourage the learners to talk with their peers about the task and discuss their ideas. This rarely leads them to copy each other, but does encourage the development of their ideas and challenges their misconceptions. These tasks are not suitable for use under test conditions; such an approach stifles the opportunities for learning.
>
> **Standard approach**
>
> Starter activity (5-10 minutes) to introduce the task. Make sure each learner knows which level they should be aiming for.
>
> Main activity (30-40 minutes) – learners attempt task. Teacher circulates, encouraging use of the level ladder and challenging misconceptions.
>
> Plenary activity (10 minutes) – self or peer assessment, where improvement ladders are used to decide on level and improvement targets.
>
> Homework activity – make the improvement, teacher collects and assesses them, giving one improvement target.
>
> **Alternative approaches**
>
> Since I wrote this section in the first edition, I have seen many teachers using these tasks in a whole variety of ways. I have seen teachers adapt the tasks to suit the needs of the class, groups of learners working together on a task, the imaginative use of ICT to respond to the task. Sometimes colleagues say, 'I hope you don't mind but I have adapted your task to…' whatever they have done. This is fantastic and I have seen some brilliant adaptations to the tasks themselves and the use of the tasks in the classroom. I would say, if you adapt the level ladders significantly, do refer back to the generic level ladder in this introduction.

> More recently, I have found more success with the tasks if I focus on the improvements instead of the first draft of the task. Once learners are confident at using the tasks, set the task as homework. Then mark this and spend the lesson focused on improvements.
>
> Identify three or four main improvement targets, find individual tasks that will challenge these (e.g. use textbook X to help you draw and label a cell diagram or complete worksheet Y to help you explain how specialised cells are adapted to their jobs). Learners choose the task related to their improvement targets. This is real AfL!
>
> (Grevatt, 2008 p. 6-7)

This led me to summarise what I perceive a pen-portrait for a teacher who uses the LATs summatively and one who uses the LATs formatively. In Table 4.2 I have summarised these characteristics. The 'Typical teacher' is mostly based on the findings of the first section of the online questionnaire. I have proposed the pen portraits of the 'Summative Teacher' and the 'Formative teacher' based on the key characteristics of the 'Typical teacher'. The statements in italics are my 'proposed' statements that characterise these teaching approaches.

Table 4.2 Pen portraits of how teachers use the LATs.

Summative Teacher	***Typical LAT teacher***	**Formative Teacher**
Uses LAT in test conditions	Uses LATs in lessons *allowing pupils to work together and use books.*	Uses LATs in lesson allowing pupils to work collaboratively
Marking provides a level (and sub level)	Marking provides a level (and sub level) and improvement target	*Marking provides a level (and sub level) and improvement target*
Occasionally allow pupils to self or peer assess to provide a level (and sub level)	Often allow pupils to self or peer assess to provide a level (and sub level) and improvement target	*Often allow pupils to self or peer assess to provide a level (and sub level) and improvement target*
Never gives pupils an opportunity to improve.	Occasionally provides pupils an opportunity to improve	*Always gives pupils an opportunity to improve*

The important outcome of this is that it has added the dimension of how the tasks were intended for use and how science teachers are actually using them.

4.1.6 Discussion

According to Bassey's (1999) educational research map, this would be an example of statistical generalisation. From this online questionnaire, I am able to make some statements, with some certainty that the LATs are used by at least 104 teachers and they are used in a variety of ways. How typical this use is of all teachers that use the task is less certain. Assuming about one thousand schools use the Badger LATs (based on sales), this represents only 10% of LAT users. The sample of teachers was opportunistic rather than a representative sample. The respondents who offered their Local Authority revealed that there is a bias to South-East England making up 50% of the total responses. However, 42 of the 150 (28%) Local Authorities in England were represented and all the regions have at least one response from a school. This is the biggest survey that has been conducted about LAT use by teachers and is consistent with my experience from working with teachers around the country.

4.2 Teachers' views on the impact on teaching and learning

This part of the questionnaire was designed to establish the views of teachers on the impact of using the LATs on teaching and learning. The quantitative analysis is displayed in Table 4.3. In the questionnaire, there was the opportunity for teachers to write specific comments relating to the LATs improving learning. These were sorted into categories, which are summarised in Table 4.4.

4.2.1 Impact on Teaching

About two-thirds of the teachers agreed or strongly agreed that their understanding of NCLDs has increased since using the tasks. With 12 stating that they do not agree with this statement. For the latter it may be reasonable to speculate that this is due to the teachers already having a good grasp of the levels before using the LATs. This assertion is supported later in the analysis.

Table 4.3 Teachers' views of the Level Assessment Tasks

Statement	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
The LATs have improved my understanding of science national curriculum levels.	12	**48**	25	12	0
The LATs have improved my teaching using Assessment for Learning strategies.	24	**49**	20	4	0
The LATs have improved the learning of my students.	10	**60**	23	4	0
My students have an improved awareness of their current level in science since using the LATs.	21	**57**	14	4	1
The LATs have hindered my teaching of science.	0	1	11	**45**	40
My students have an improved awareness of how to improve in science, since using LATs.	9	**61**	22	5	0
The LATs have hindered the learning of my students.	0	0	5	42	**50**
The LATs have improved my assessment of my students.	14	**60**	20	3	0
My students enjoy doing the LATs.	9	**49**	27	8	4
I enjoy teaching using the LATs.	11	**58**	24	3	1

N= 97, 7 respondents skipped this section. ***Bold numbers*** are highest response number.

A larger majority of the respondents (75%) felt that their AfL strategies had improved since using the tasks. However, this does not tally with the finding that most of these teachers do not regularly give their pupils opportunity to improve. So the question that arises from this is what aspects of AfL do they believe have improved?

A possible answer to this can be found in some of the additional comments that the teachers who responded provided (Table 4.4). Teachers cited the identification of misconceptions as a specific way in which the tasks support teaching. This may be an interpretation of supporting AfL, which could happen without teachers setting specific improvement targets. Examples of comments included general comments about identifying misconceptions:

Better understanding of poorly understood areas and miscomprehensions.

I have been able to assess areas that I need to readdress.

And a specific misconception that a teacher identified:

They have helped me to identify student misconceptions eg many students labeled the whole plant 'flower'.

88% of teachers do not feel that the use of the tasks have hindered their teaching, however, more interesting eleven were neutral and one agreed with the statement. This causes me to consider what is meant by ‘neutral’? Does that mean the teacher has no opinion or that they think the tasks have neither helped nor hindered their teaching? The one teacher who does believe the tasks have hindered their teaching could, optimistically, be an input error or if that is what they believe, it would be interesting to know in what ways, something beyond the scope of this questionnaire. The additional comments provide some insight. Some of the teachers that felt that using the LATs had made no difference to their teaching offered these statements:

I don't feel it has made a difference to my teaching style

I wouldn't say they have improved my teaching, they are one of a variety of tasks I use with my groups but if they didn't exist I would use one of the similar activities I have.

Doing similar before so has improved access to material and not necessarily improved my teaching. Just added additional resources.

The general feeling appears to be that these teachers already had the LAT approach in place and that the tasks supported what was already being done.

With respect to assessment, over three-quarters of teachers agree that the LATs have improved their assessment of the pupils. The next question is therefore, how have the tasks helped? With respect to the 20 ‘neutral’ and 3 ‘disagree’ responses it is possible

that the teachers felt perfectly adequate at assessing their pupils as they did above. I wonder if the neutral statements mean that that the teachers have not observed an improvement in their assessment because they are using the tasks summatively.

Table 4.4 Summary of qualitative responses to LATs supporting teaching.

To what extent has using the LATs improved your teaching?	Number of responses
Increased 'Level Awareness'	22
Identifying misconceptions	11
Supports AfL	10
Supports progression	6
No difference	5
Using self and peer assessment	4
Consolidates learning	3
Improved assessment	3
Supports differentiation	2
Increased expectation	2

In the qualitative comments (Table 4.4), the view that using LATs had improved their 'level awareness' was the predominant comment. Representative statements included:

> *I am more aware of the answers required for different levels. This has helped when setting learning objectives at the start of lessons.*
>
> *Made me more aware of specifics of NC levels - also, even in lessons where I wasn't using the tasks, it made me think a lot more about levels I needed to pitch lessons at and to whom differentiation should be addressed.*
>
> *Improved my breadth of understanding of what students should consider.*
>
> *Made me and the staff in my department much more level aware. It has increased the amount of self and peer assessment of the students which in turn has increased motivation.*

The last comment specific includes the term 'level aware' which I think encapsulates the idea of the tasks supporting teachers (and pupils, it will be demonstrated later) with engaging with the notion of 'levelness'.

Ten teachers stated specifically that the tasks had helped improve their understanding or use of AfL. Their comments included:

> *Allowed me to use Assessment for Learning - something I previously struggled with as a trainee teacher.*
>
> *Increased the amount of formative assessment I use.*
>
> *It has made Assessment for Learning more consistent throughout our KS3 topics.*

However I would raise the question again, is AfL really happening when teachers are using these tasks? I would argue that teachers feel that by setting a target, as characterized by the pen portraits in Table 4.2, is not in itself AfL. However, teachers may be referring to being able to identify their pupils' misconceptions and then challenge these as a group, rather than on an individual basis.

4.2.2 Impact on learning

In Table 4.3, seventy of the teachers agreed or strongly agreed that the LATs had improved the learning of their pupils, with 4 disagreeing with the statement. In contrast, no teachers agreed that the use of the LATs hindered their pupils' learning, with half the teachers strongly disagreeing with this statement. In addition a large number of the teachers agreed that using the tasks had improved their pupils' understanding of the NCLDs.

The coded qualitative data in Table 4.5 shows how teachers perceived that the LATs support their pupils' learning. With respect to *how* teachers think that the LATs have improved learning, these data provides some insight. The most common ways being: pupils' understanding of the requirements of each level, being able to identify their own improvements, the LATs supporting pupil progression and improved communication in science.

Table 4.5 Summary of qualitative responses to pupils' learning from using LATs

To what extent has using the LATs improved the learning of your students?	**Number of responses**
Understanding the requirements of each level	19
Identify Improvements	13
Supports Progression	9
Improved communication is science	8
Supports independent learning	6
Improved use of science	5
Awareness of own level	5
Demotivating	2
Aids peer and self assessment	2
Supports the more able	1
Supports the less able	1
Other	9

Pupil's improved understanding of 'levels' was cited as:

> *Students understand more what is required for certain levels. They are having to explain themselves more*
>
> *It has helped me to focus pupils on the higher level criteria that they need to move up a level, e.g. word equations and using chemical symbols, which has helped pupils to know what they need to do to reach a higher level.*
>
> *My students understand their levels better and are able to better apply their science knowledge and understand how they can improve.*
>
> *They can clearly see what is required to achieve each level and how to improve*

This is related to, but different to the ability of pupils to identify their own areas for improvements. Some teachers recognised this as something that the LATs have supported:

> *Students are aware of their level of knowledge and understanding and can be targeted to improve on certain areas of weakness. The use of peer and self assessment has helped my students judge their own level of knowledge/skills and allowed them to set their own target.*
>
> *They know exactly what is expected of them. They can assess what they have done and what else they need to do by using the level ladder…*
>
> *Allowed students to determine what level they are working at, and what they need to do to improve their level with specific examples.*

The improved communication of science of their pupils was cited by some teachers:

> *They are improving their communication of science. The open tasks are giving them more space to show what they can do.*
>
> *Pupils better able to apply a model to explain phenomena*
>
> *they are much more used to writing SCIENCE not waffle*
>
> *Ability to link ideas and consolidate Use of keywords from the list*
>
> *More focus to key words and overall understanding of a concept*

This, I contend is in contrast to the often trivial aspects of learning that pupils associate with making 'improvements' such as 'write more neatly' or 'colour in my title'.

4.2.3 Summary of the impact of using the LATs on teaching and learning

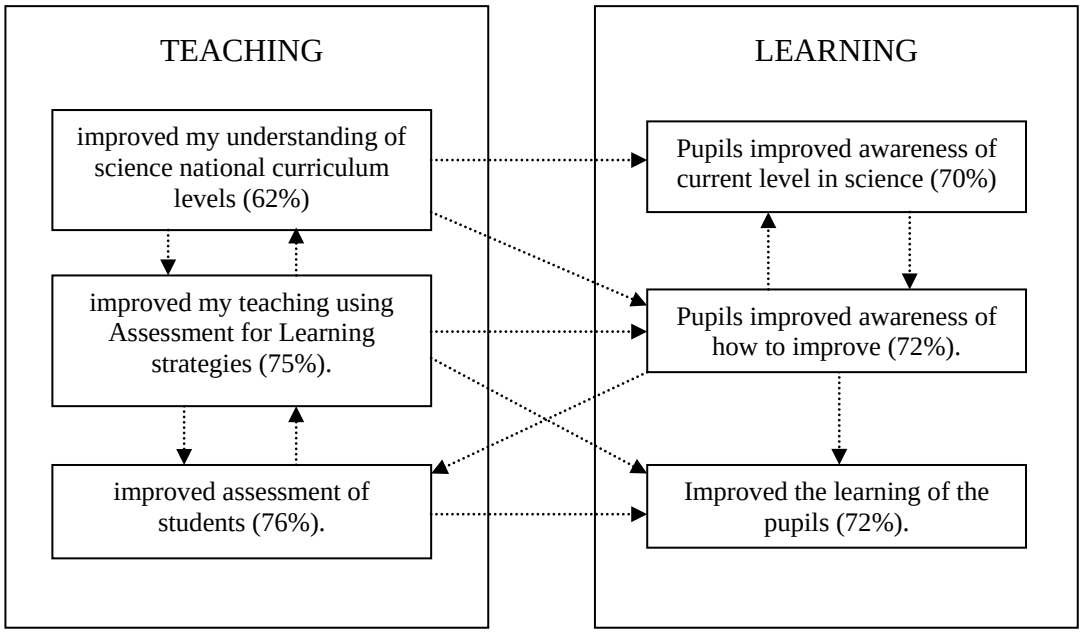
In order to conceptualise the main views expressed by the teachers who use the LATs, I have put the main features of the analysis into a descriptive framework. Figure 4.1 is my attempt to summarise and describe the impact of the LATs on teaching and learning based on the responses of teachers to the online questionnaire. The statements with which a majority of teachers agreed or strongly agreed have been put into the framework.

I then considered how these might be related and used dotted arrows to indicate relationships between the statements. These arrows could be conceived as 'fuzzy relationships' (Bassey, 1999). They are how I propose that the elements could develop and are related.

The 'teaching box' is composed of the three statements associated with assessment based on levels (taken from Table 4.3). A majority of the teachers were of the opinion that using the LATs had improved their understanding of the science NCLDs, improved their use of AfL strategies and improved their assessment of their students. I then

considered how these three elements might be related and decided that a reasonable proposal would be that when using the LATs, a teacher's understanding and confidence in the NCLDs might improve. This improvement would lead to teachers' ability to understand at which level their pupils are working and what they need to do to improve, thus improving their AfL strategies. It may be worth noting at this point that uncovering and challenging misconceptions could also be considered as an AfL strategy. Improved AfL strategies could then lead to a teacher being able to better assess their students.

Figure 4.1 A framework to describe the impact of the tasks on teaching and learning



The arrows between these boxes are intended to suggest a relationship between these elements. For example, as the teacher's use of AfL strategies improves, they may refine their understanding of the NCLDs as they try these strategies in the classroom, thus further improving their understanding of levels.

The 'learning box' has the three elements that a majority of teachers agreed had improved through using LATs, based on Table 4.1. These elements were pupils'

improved awareness of their current level in science, improved understanding of how to improve their work and improved learning from using the LATs. Again, I considered how these elements might be related. The sequence I propose is that as a pupil has an improved awareness of their current level, they will be able to perceive a gap in their learning and identify what they need to do to improve and once this has been established the pupil may improve their understanding of science.

The arrows between the boxes are again intended to suggest a relationship between the statements. Specifically, I suggest that a pupil will refine their understanding of their current level as they learn to appreciate the gaps in their knowledge, so there are arrows going both ways between these boxes. This can lead to improved learning, but I did not feel that there would be a connection from improved learning to either of the two other boxes.

I then considered possible relationships between the teaching elements and the learning elements that I had identified. It may be reasonable to suggest that there is a relationship between the teachers' improved understanding of levels and the pupils' level awareness and the pupils' improved understanding of how to improve. For example, as the teacher improves their understanding of levels, their confidence to discuss them with pupils will increase and lead to conversations between them about where they are and how to improve. The teachers' improved use of AfL strategies is likely to lead to the pupils' understanding of how to improve and thus their learning in science. This type of relationship is well established through educational research of formative assessment strategies explored in Black and Wiliam (1998b). I also propose that when pupils have a better awareness of how to improve, this will enhance the teacher's assessment of their pupils. This is because once a dialogue is established teachers may gain further insight into the understanding of their pupils.

4.2.4 Summary of findings from questionnaire

From this analysis the following statements can be made from this sample of teachers:

To what extent do the LATs support teaching?

- The LATs *can* improve teachers' understanding of science NCLDs
- The LATs *can* improve teachers' use of AfL strategies
- The LATs *can* improve teacher assessment of students

To what extent do the LATs support learning?

- The LATs *can* improve pupils' awareness of their current level in science
- The LATs *can* improve pupils' awareness of how to improve
- The LATs *can* improve pupils' learning in science.

4.3 Specific research questions

From this questionnaire, more specific research questions could be raised.

To what extent do the LATs support teaching?

- Do teachers use the tasks in the way they say they do?
- Why do some teachers not give pupils the opportunity to improve?
- What are the differences between the 'summative', 'typical' and 'formative' teacher when using the LATs?

To what extent do the LATs support learning?

- What is the pupil perspective? How do pupils feel about the tasks?
- How do they LATs improve pupil awareness of levels and improvement?

What features of LAT use makes them formative?

These questions are explored by the case studies of the five lessons described below.

4.4 Case studies

I have carried out five case studies of the use of a task in a classroom by different teachers. Each case study has five components:

- Lesson observation
- Background data
- Pupil questionnaire
- Pupil interview (group of 3-5)
- Teacher interview

This generated a large amount of data, too much to present within the thesis. Instead I present the relevant and pertinent data that can give the reader a logical and detailed summary of the data that have led to my conclusions.

4.4.1 Summary of the case studies

The five cases studies took place in four different schools, by five different teachers (Table 4.6). All Key Stage 3 year groups are represented and the experience of the teachers using the tasks ranged from the first time (Summative Case) to 4 years (Typical and Transitional Cases). The cases are named after the assessment 'attitude' approach I

thought the teacher was using to deliver the LAT for the case study. It should be noted that I am assuming that the claims I make are about any LAT-style activity, not the specific task used for a particular case study.

Table 4.6 Case Study Summaries

CASE:	**Formative**	**Mismatch**	**Summative**	**Typical**	**Transitional**
School	S1	S1	S2	S3	S4
Teacher	T1	T2	T5	T8	T9
Year	8	9	9	7	8
Level Assessed Task	8K Energy Transfers	9A Healthy Living	9K Gravity and Space	7C Food webs	8F Burning Magnesium
Pupils Male/female/Total	12/11/23	13/12/25	10/13/23	12/15/27	7/11/18
Ability range (Levels)	6-7	6-7	3-5	3-5	3-5
Time using tasks	3 years	3 years	first time	4 years	4 years
Pupils interviewed (pseudonyms)	Teagan Isobel Mac	Toby Mike Rose Rachel	Imogen Jessica Kyle Liam	Amanda John Mark	Andy Brian Cathy Dawn

The specific LAT used in each Case Study are displayed in Figures 4.2-4.6:

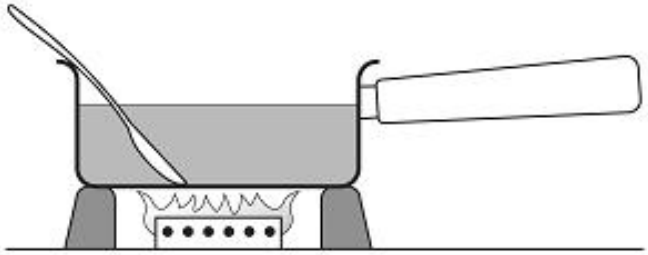
Figure 4.2 Level Assessed Task used in the Formative Case Study

8I Task Sheet (L5-7)

Heat in the kitchen

During a Food Technology lesson, some students were wondering why the metal spoon gets hot, but the wooden handle of the saucepan does not.
Use your knowledge and understanding to explain how the energy is transferred from the cooker to the end of the metal spoon.

Task:
Show how the energy from the gas ring is transferred to the end of the spoon.
Describe how the energy is transferred.
Use particle diagrams to show what is happening.



Key words: conduction, conductor, convection, energy transfer, evaporation, heating, insulator, radiation

Level ladder: What is your target level? Use the level ladder to help you reach it:

To get level	You might have:
5	• Drawn or used a diagram of the saucepan and spoon. • Used the key words metal, wood, water, insulator and conductor to label the diagram. • Drawn arrows to show heat energy transfers from the gas ring to the spoon handle. • Stated why energy is transferred from the hob to the spoon. • Identified energy transfers by conduction, convection and radiation. • Used a simple particle diagram to explain conduction.
6	• Drawn or used a diagram of the saucepan and spoon. • Described the diagram using the key words metal, wood, water, insulator and conductor. • Used simple energy transfer diagrams to show heat energy transfers from the gas ring to the spoon handle. • Explained why energy is transferred from the hob to the spoon. • Used particle diagrams to explain conduction and convection.
7	• Drawn or used a diagram of the saucepan and spoon. • Described the diagram using the key words metal, wood, water, insulator and conductor. • Used particle diagrams and energy transfer diagrams to show heat energy transfers from the gas ring to the spoon handle. • Explained in detail why energy is transferred from the hob to the spoon. • Used detailed scientific knowledge and understanding to compare energy transfer by conduction, convection, radiation and evaporation.

Figure 4.3 Level Assessed Task used in the Mismatch Case Study

9B Task Sheet (L5-7)

Healthy lifestyle?

When you go to a new doctor's surgery, you have to fill in a form about your lifestyle. Read what this male patient has written.

Question	Patient
Do you smoke?	*Yes*
If so, how much?	*10 a day*
Do you drink alcohol?	*Yes*
If so, how often and how much?	*A few cans of larger in the evening and spirits at the weekend*
How often do you exercise?	*Once a week*
What do you do?	*1 hour at the gym*
What types of food do you snack on?	*Crisps, chocolate bars*
How often do you eat fresh fruit and vegetables?	*Eat potatoes with most meals, usually as chips*

Task: From this information, use your scientific knowledge and understanding to:

- Identify what is healthy about the patient's lifestyle.
- Predict health problems the patient may have from their lifestyle.
- Suggest some changes that they could make to their lifestyle to make it healthier.
- Explain the short-term and long-term effects of the lifestyle on their body.

Key words: addiction, alveolus, arteries, bad breath, blood, cancer, cholesterol, cilia, emphysema, fitness, heart, heart attack, lungs

Level ladder: What is your target level? Use the level ladder to help you reach it:

To get level	You might have:
5	• Identified some healthy and unhealthy parts of the patient's lifestyle. • Suggested some changes the patient could make to be healthier, stating a scientific reason. • Stated the main organs or **body systems** and **organs** it can effect. • Described some health problems with smoking, drinking too much alcohol and eating too many fatty foods. • Described the effects of exercise on the body.
6	• Described the healthy and unhealthy parts of the patient's lifestyle. • Suggested a range of changes the patient could make to be healthier, using scientific knowledge. • Explained the effects on the main **body systems**, **organs** and **cells**. • Explained a range of short-term and long-term health problems with smoking, drinking too much alcohol and eating too many fatty foods. • Explained the effects of exercise on the body.
7	• Explained the healthy and unhealthy parts of the patient's lifestyle. • Suggested a range of changes the patient could make to be healthier, using detailed scientific knowledge. • Explained the effects on the main **body systems**, **organs** and **specialised cells**. • Explained a range of short-term and long-term health problems with smoking, drinking too much alcohol and eating too many fatty foods. • Explained the positive and negative effects of exercise on the body.

Figure 4.4 Level Assessed Task used in the Summative Case Study

9J Task Sheet (L3-5)

Interplanetary postcards

Our interplanetary scientist, Cefor, has sent some electronic cards from her visits to places in our solar system. You have asked her to tell you her weight and mass at each place.

When Cefor left Earth, she had a mass of 10kg and a weight of 100N.

Task:
Draw a force diagram for Cefor standing still on Earth and on the three other places. Use the level ladder instructions to help explain your ideas.

Key words: acceleration, force, gravity, kilograms, mass, Newtons, reaction

Level ladder:
What is your target level? Use the level ladder to help you reach it:

To get level	You might have:
3	*Stated simply why Cefor's weight is different at each place by:* • Drawing a diagram of Cefor on each place. • Identifying where there may be a force.
4	*Described how Cefor's weight is different at each place by:* • Drawing a diagram of Cefor in each place, using some force arrows. • Labelling arrows with key words. • Stating whether the forces are balanced or unbalanced. • Defining simply "mass" and "weight".
5	*Explained simply why Cefor's weight is different at each place by:* • Drawing a diagram of Cefor in each place, using force arrows that show the direction of the force, and labelling arrows with key words. • Stating whether the forces are balanced or unbalanced. • Describing the difference between "mass" and "weight". • Labelling the size of the forces on each diagram.

Figure 4.5 Level Assessed Task used in the Typical Case Study

7C Task Sheet (L5-7)

The rock pool food web

An ecologist has been studying a rock pool for a week. She has noted which plants and animals she has seen and what they are eating. Below are her notes.

Rock Pool Feeding Relationships: observation notes

Monday: *I saw a* ***shore crab*** *eating some* ***green seaweed****. A* ***sea anemone*** *was eating by catching small particles of floating plants and animals.*



Tuesday: ***Whelks*** *are eating* ***red*** *and* ***green seaweed****.* ***Small fish*** *dart out from under rocks and catch* ***shrimps****.*



Wednesday: *The* ***shrimps*** *seem to be eating all the time. They are eating microscopic plants and animals. These organisms are called* ***plankton****.* ***Animal plankton*** *eats* ***plant plankton*** *in the rock pool. It is likely that* ***sea anemones*** *are eating* ***plankton*** *too.*


Thursday: *Observed the* ***shore crabs*** *more closely today. One was eating a* ***shrimp*** *it had caught. Another was eating a* ***small fish****, but I did not see it catch it. The* ***crabs*** *also eat* ***whelks****.*

Friday: ***Small fish*** *nibble at* ***sea anemones****.* ***Seagulls*** *occasionally land next to the pool. I have seen them eat* ***small fish*** *and* ***crabs****.*

Task: Write down some daily and seasonal changes that might happen in the rock pool habitat. Choose a predator and prey and write down some adaptations that help them to survive. Draw some food chains or a food web using some or all of this information. Decide which type of feeder each organism is (primary consumer, etc) and identify the producers.

Level ladder: What is your target level? Use the level ladder to help you reach it:

To get level	You might have:
5	• Drawn a simple food web that links at least three food chains. • Used some key words to describe how plants and animals get their food (e.g. producer, primary consumer, carnivore). • Described at least one adaptation that helps a predator and prey to survive. • Described one daily **and** one seasonal change that affects the rock pool habitat.
6	• Drawn a simple food web that links most of the organisms. • Used key words to describe how plants and animals get their food (e.g. producer, primary consumer, carnivore). • Explained at least one adaptation that helps a predator and prey to survive. • Explained how daily **and** seasonal changes affect the rock pool organisms. • Used the Big Idea of Energy to explain how energy flows through the food chains in the rock pool.
7	• Drawn a detailed and organised food web, with all types of feeders correctly identified. • Explained several adaptations that help a predator and prey to survive. • Explained daily **and** seasonal changes that affect the rock pool organisms. • Used the Big Idea of Energy to explain how energy flows through the food chains in the rock pool.

Figure 4.6 Level Assessed Task used in the Transitional Case Study

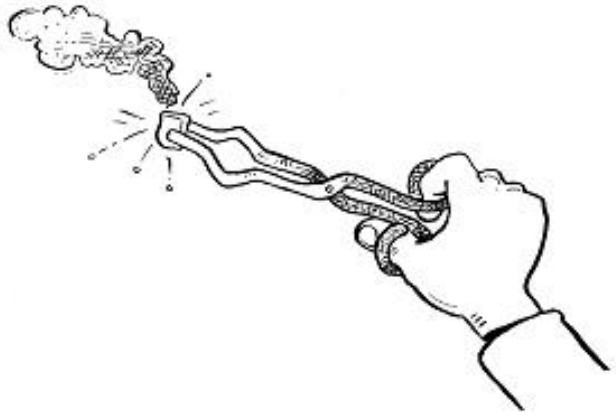
8E Task Sheet (L5-7)

Explaining what happens when we burn magnesium metal

Watch the reaction when magnesium burns in air.

Task:
Explain what happens when magnesium burns.

Top Tips:
- Try to use particle diagrams to show what is happening.



Key words: compound, element, metal, mixture, non-metal, reaction

Level ladder:
What is your target level? Use the level ladder to help you reach it:

To get level	You might have:
5	• Identified the elements, compounds or mixtures. • Described the appearance and properties of the materials before and after the reaction (classified as metals or non-metals). • Described what happens during the reaction. • Used a simple particle model to describe the change.
6	• Used a particle model to explain the reaction in detail. • Described why the substances can be classified as an element, compound or mixture. • Explained the reaction using the Big Idea of Energy. • Explained why there are the same numbers of particles in the beginning and at the end of the reaction. • Written the word equation for the reaction.
7	*Followed the instructions for level 6, using detailed scientific knowledge and understanding, and also:* • Used symbols from the Periodic Table to write a balanced symbol equation.

4.5 Analysis of the case study pupil questionnaire.

At the end of each lesson, pupils were given a short questionnaire to complete about the task they had just carried out. This data is presented in Tables 4.7 and 4.8, Questions 1-7. Table 4.7 summarises the unadjusted aggregate of all five case studies. Tables 4.8, questions 1-7 display the adjusted aggregated data to compare each case study. The raw data and the percentage response within the class are presented, as well as an aggregate of all the case studies combined. I have used the adjusted aggregate as a comparison baseline. For each question, in the case of any one of the five cases, the aggregation of the data of the other four cases is quoted (% aggregate) for comparison with that case's result (% within case). The complete spreadsheet is shown in Appendix G. The question of gender differences arose from this data (Appendix H). However, a statistical analysis using chi-squared, showed that there was no significant difference between male and female responses for any question (Appendix I).

Table 4.7 Aggregate Pupil Post-Lesson Questionnaire

		Q1		Q2		Q3		Q4		Q5		Q6		Q7	
	Sex	raw	%	raw	%	raw	%	raw	%	Raw	%	raw	%	raw	%
Strongly Agree	M	17	31	10	19	1	2	9	17	25	46	1	2	1	2
	F	8	14	17	29	1	2	10	17	25	42	7	12	12	20
	T	**25**	**22**	**27**	**23.9**	**2**	**1.8**	**19**	**16.8**	**50**	**44.2**	**8**	**7.1**	**13**	**11.5**
Agree	M	26	48	33	61	3	6	29	54	17	31	24	44	12	22
	F	28	47	33	56	4	7	32	54	27	46	28	47	15	25
	T	**54**	**47.8**	**66**	**58.4**	**7**	**6.2**	**61**	**54.0**	**44**	**38.9**	**52**	**46.0**	**27**	**23.9**
Not Sure	M	11	20	1	2	15	28	8	15	9	17	22	41	10	19
	F	17	29	4	7	19	32	9	15	3	5	12	20	13	22
	T	**28**	**24.8**	**5**	**4.4**	**34**	**30.1**	**17**	**15.0**	**12**	**10.6**	**34**	**30.1**	**23**	**20.4**
Disagree	M	1	2	8	15	32	7	5	9	3	6	4	7	11	20
	F	4	7	3	5	27	12	4	7	0	0	7	12	13	22
	T	**5**	**4.4**	**11**	**9.7**	**59**	**52.2**	**9**	**8.0**	**3**	**2.7**	**11**	**9.7**	**24**	**21.2**
Strongly Disagree	M	0	0	3	6	4	7	0	0	1	2	2	4	21	39
	F	2	3	2	3	7	12	2	3	3	5	4	7	6	10
	T	**2**	**1.8**	**5**	**4.4**	**11**	**9.7**	**2**	**1.8**	**4**	**3.5**	**6**	**5.3**	**27**	**23.9**

M=Male (n=54), F=Female (n=59), T=Total (n=113)

Key :
Q1 : I enjoyed doing the task
Q2 : I used the Level Ladder to help me do the task.
Q3 : I found this task difficult to do.
Q4 : The Level Ladder showed me how to improve.
Q5 : I learnt something new
Q6 : The Level Ladder helped me instead of asking my teacher
Q7 : Another pupil helped me when I got stuck

Table 4.8 Pupil Post-Lesson Questionnaire: Question 1 Analysis

Question 1: I enjoyed doing the task

Q1		Formative Case Study	Mismatch Case Study	Summative Case Study	Typical Case Study	Transitional Case Study
Strongly Agree	Raw	7	3	6	1	3
	% within case	30	12	26	4	17
	% aggregate*	14	19	15	21	17
Agree	Raw	13	20	8	12	3
	% within case	57	80	35	44	17
	% aggregate*	46	40	52	49	54
Not Sure	Raw	3	2	6	12	7
	% within case	13	8	26	44	39
	% aggregate*	29	31	26	20	23
Disagree	Raw	0	0	1	2	3
	% within case	0	0	4	7	17
	% aggregate*	0	0	5	4	3
Strongly Disagree	Raw	0	0	2	0	2
	% within case	0	0	9	0	11
	% aggregate*	0	0	2	0	2

Table 4.8 Pupil Post-Lesson Questionnaire: Question 2 Analysis

Question 2: I used the Level Ladder to help me do the task

Q2		Formative Case Study	Mismatch Case Study	Summative Case Study	Typical Case Study	Transitional Case Study
Strongly Agree	Raw	5	1	4	12	0
	% within case	22	4	17	44	0
	% aggregate*	18	23	19	11	0
Agree	Raw	16	18	11	8	10
	% within case	70	72	48	30	56
	% aggregate*	51	49	56	62	54
Not Sure	Raw	0	5	6	4	1
	% within case	0	20	26	15	6
	% aggregate*	17	12	11	13	15
Disagree	Raw	2	0	2	2	3
	% within case	9	0	9	7	17
	% aggregate*	8	0	8	8	6
Strongly Disagree	Raw	0	1	0	1	4
	% within case	0	4	0	4	22
	% aggregate*	0	3	0	3	0

Table 4.8 Pupil Post-Lesson Questionnaire: Question 3 Analysis

Question 3: I found this task difficult to do

Q3		Formative Case Study	Mismatch Case Study	Summative Case Study	Typical Case Study	Transitional Case Study
Strongly Agree	Raw	0	0	1	0	2
	% within case	0	0	4	0	11
	% aggregate*	0	0	2	0	1
Agree	Raw	2	0	3	1	0
	% within case	9	0	13	4	0
	% aggregate*	4	0	3	6	0
Not Sure	Raw	5	4	6	11	8
	% within case	22	17	26	42	44
	% aggregate*	32	33	31	26	27
Disagree	Raw	15	17	8	11	3
	% within case	65	71	35	42	17
	% aggregate*	43	41	51	49	53
Strongly Disagree	Raw	1	3	5	3	5
	% within case	4	13	22	12	28
	% aggregate*	18	16	13	16	13

Table 4.8 Pupil Post-Lesson Questionnaire: Question 4 Analysis

Question 4: The Level Ladder showed me how to improve

Q4		Formative Case Study	Mismatch Case Study	Summative Case Study	Typical Case Study	Transitional Case Study
Strongly Agree	Raw	3	5	4	6	4
	% within case	14	20	17	24	24
	% aggregate*	21	20	20	18	19
Agree	Raw	15	11	11	11	5
	% within case	68	44	48	44	29
	% aggregate*	42	48	47	48	51
Not Sure	Raw	2	7	6	7	4
	% within case	9	28	26	28	24
	% aggregate*	27	22	22	22	23
Disagree	Raw	2	1	1	1	2
	% within case	9	4	4	4	12
	% aggregate*	6	7	7	7	5
Strongly Disagree	Raw	0	1	1	0	2
	% within case	0	4	4	0	12
	% aggregate*	0	3	3	0	2

Table 4.8 Pupil Post-Lesson Questionnaire: Question 5 Analysis

Question 5: I learnt something new

Q5		Formative Case Study	Mismatch Case Study	Summative Case Study	Typical Case Study	Transitional Case Study
Strongly Agree	Raw	11	3	1	11	6
	% within case	48	14	5	42	33
	% aggregate*	24	33	35	25	28
Agree	Raw	10	13	16	11	3
	% within case	43	62	73	42	17
	% aggregate*	49	45	42	50	54
Not Sure	Raw	2	3	2	1	5
	% within case	9	14	9	4	28
	% aggregate*	13	11	13	14	9
Disagree	Raw	0	1	1	1	2
	% within case	0	5	5	4	11
	% aggregate*	0	4	5	5	3
Strongly Disagree	Raw	0	1	2	2	2
	% within case	0	5	9	8	11
	% aggregate*	0	7	6	6	5

Table 4.8 Pupil Post-Lesson Questionnaire: Question 6 Analysis

Question 6: The Level Ladder helped me instead of asking my teacher

Q6		Formative Case Study	Mismatch Case Study	Summative Case Study	Typical Case Study	Transitional Case Study
Strongly Agree	Raw	1	1	1	3	2
	% within case	4	4	4	13	11
	% aggregate*	8	8	8	6	6
Agree	Raw	13	6	3	6	7
	% within case	57	25	13	25	39
	% aggregate*	25	33	36	33	30
Not Sure	Raw	8	8	11	9	1
	% within case	35	33	48	38	6
	% aggregate*	33	33	29	32	38
Disagree	Raw	1	7	6	4	4
	% within case	4	29	26	17	22
	% aggregate*	24	17	18	20	19
Strongly Disagree	Raw	0	2	2	2	4
	% within case	0	8	9	8	22
	% aggregate*	0	9	9	9	6

Table 4.8 Pupil Post-Lesson Questionnaire: Question 7 Analysis

Question 7: Another pupil helped me when I got stuck

Q7		Formative Case Study	Mismatch Case Study	Summative Case Study	Typical Case Study	Transitional Case Study
Strongly Agree	Raw	1	2	3	6	4
	% within case	4	8	13	22	22
	% aggregate*	16	15	14	11	12
Agree	Raw	**6**	**13**	**5**	**5**	**4**
	% within case	26	52	22	19	22
	% aggregate*	29	22	30	31	30
Not Sure	Raw	**7**	**4**	**1**	**1**	**1**
	% within case	30	16	4	4	6
	% aggregate*	8	11	14	15	13
Disagree	Raw	**6**	**5**	**12**	**4**	**2**
	% within case	26	20	52	15	11
	% aggregate*	25	26	18	28	28
Strongly Disagree	Raw	3	1	2	11	7
	% within case	13	4	9	41	39
	% aggregate*	23	25	24	15	17

* For aggregate calculation – see Appendix G

The pupil questionnaires were my main source of information about pupils' attitudes to the LATs. I will use this as part of triangulation in the analysis in Chapter 5. For now, it is useful to make some generalisations from these 116 pupils' views. In this instance, I will refer only to the aggregate data in Table 4.7.

With regard to the research questions of the extent to which the tasks support learning, it could be argued that if pupils enjoy what they are doing, they are more likely to learn. In direct response to the question of whether they learnt something new while doing the task (Q5), 74% of pupils believed that they learnt something new, 14% did not. What supported that learning could be the Level Ladder on the task, the teacher or another pupil.

A vast majority of pupils claimed that they used the Level Ladder to help them do the LAT (73%), with 13% claiming they did not (Q2). In comparison 65% claimed that the Level Ladder helped them improve (Q4), with 9% stating that it did not. Question 7 probed to find if pupils used the Level Ladder for support instead of the teacher, to which only 37% of pupils agreed or strongly agreed, with 28% disagreeing or strongly disagreeing. The most divisive question regarded the use of another pupil to help with the task where 42% claimed to use another pupil, but 46% did not. This would suggest that some peer learning is taking place, but is not seen consistently by pupils as something they do while doing a LAT.

Of the pupils who completed the questionnaire, 65% stated that they enjoyed doing the LAT, 26% were unsure if they enjoyed the task and a further 8% did not enjoy doing the task. This may be linked to how difficult the pupils found the task, where 63% claimed that they did not find the LAT hard, with 8% claiming that they found the LAT difficult to do. Obviously the pupils had varied experiences of the delivery of the LAT,

so how generalisable these data are, is debateable. However, it does raise a question of whether there is a relationship between how easy the pupil found the task and how much they enjoyed doing it.

The qualitative element of the pupil questionnaire was categorised based on what the pupils regarded as what they 'liked best' about doing the task (Table 4.9) and what they 'liked least' (Table 4.10). Some pupils offered only a one word answer, others a short phrase or sentence. Due to this I did not code this data in any detail, but instead just categorised the responses. The main aspect to note here is that the pupils were free to write whatever they liked for these two questions. Overall, the most common categories for what pupils liked best were learning something, the 'topic' (the subject that the LAT was about) and collaboration (being able to work together). There is variability between each case which will be discussed in the final analysis. Table 4.10 displays what pupils liked least about doing the task, which was consistent in each case and overall for 'writing.' This I took to mean having to write to communicate their ideas. The task itself was commonly cited, and mentioned at least once in each case. On their own, these data do not give a lot of insight for the research questions. It could be surmised that pupils in some cases like working together, learning something about the topic, and doing some practical work (where the opportunity arose) and liked having to write the least, something that is common in most classroom activities and not unique to LATs.

Table 4.9 Pupil Questionnaire – 'What I liked best' Qualitative categories

	Case Study					
Liked best	*Formative*	*Mismatch*	*Summative*	*Typical*	*Transitional*	*All*
Creative	1	0	0	0	0	1
collaboration	2	9	0	0	0	11
drawing	0	0	0	0	0	0
enjoyment	0	0	0	2	0	2
independent	3	0	0	0	1	4
learning something	3	2	6	7	1	19
no comment	0	0	0	0	3	3
Nothing	0	0	2	0	1	3
Other	4	4	1	3	2	14
practical	0	0	0	5	3	8
Topic	4	5	0	7	0	16
Writing	2	1	0	1	0	4
Levels	1	0	0	0	0	1

Table 4.10 Pupil Questionnaire – 'What I liked least' Qualitative categories

	Case Study					
Liked least	*Formative*	*Mismatch*	*Summative*	*Typical*	*Transitional*	*All*
Boring	0	0	1	2	0	3
Concept	0	0	3	0	0	3
Drawing	1	0	0	1	0	2
Everything	0	0	1	0	0	1
Improving	0	0	0	0	0	0
Nothing	0	3	1	4	0	8
Other	3	4	1	2	3	13
Practical	0	0	0	1	0	1
Research	1	1	0	0	0	2
Task	2	1	4	3	2	12
Time	2	1	0	1	0	4
Topic	0	0	2	1	0	3
Writing	5	8	6	12	1	32

Table 4.11 Teacher Interview Themes

Themes	*Formative*	*Mismatch*	*Summative*	*Typical*	*Transitional*
Collaborative	✓	✓		✓	
Copying (pupils each other)	✓				
Learning from task	✓	✓			
Target setting					✓
Task use		✓	✓		✓
Tests vs LATs		✓			
Time		✓			
Values knowledge and understanding	✓	✓	✓	✓	✓
Values presentation	✓		✓	✓	✓
Other	✓				

Table 4.12 Pupil Interview Themes

Themes	*Formative*	*Mismatch*	*Summative*	*Typical*	*Transitional*
Collaboration	✓	✓		✓	
Task use	✓	✓	✓		✓
Test vs LAT		✓			
Values knowledge and understanding	✓	✓	✓	✓	✓
Values presentation	✓		✓	✓	✓
Other	copying	Time			Target setting

4.6 The concepts of ‘Levels’ and ‘Levelness’.

It is timely to define the conceptual codes of ‘levels’ and ‘levelness.’ Originally the latter I defined as ‘levelling’, though I felt it was not distinct enough from ‘levels’. The concept of ‘levelness’ has been used in prep room parlance, I am unsure of its origin, but here I will define it as it forms an important thread of my thesis. Teachers and pupils who have a notion of ‘levels’ treat the level descriptors on the Level Ladder literally and use the levels as numbers, marks or grades. In contrast, with a notion of ‘levelness,’ teachers and pupils have an understanding of the generic requirements of a particular level. This means that they have an understanding of the concept of levels and how they relate to one another, possibly supported by the Levels Mountain or similar described in

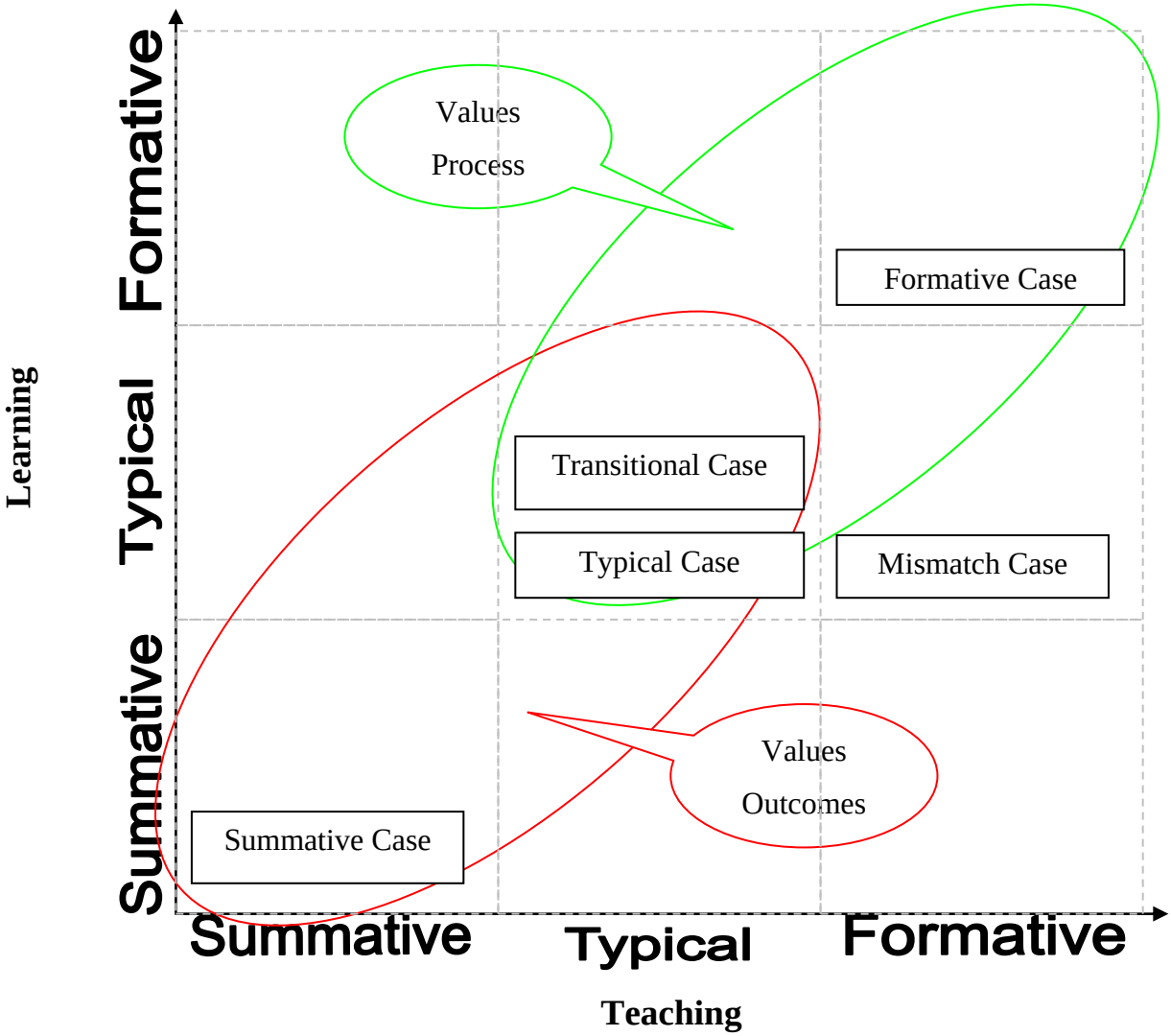
Chapter 1. The concepts of ‘levels’ and ‘levelness’ will be explored further in the subsequent chapters.

Admittedly there is a lot of data presented here. It is important to present this as it forms the backdrop for the analysis in subsequent chapters. The last stage of the coding of the interviews is axial coding, where relationships between the categories are considered. This forms the main analysis in Chapter 5.

Chapter 5 Analysis and Fuzzy Generalisations

The purpose of this research was to explore *how* the LATs were used and explain when the LATs best support teaching and learning. As a result, I was able to identify some characteristics of teaching approaches, which I exemplified in 'pen portraits' of teachers (Table 4.2). In Figure 5.1 I present a hypothesis via a graph onto which I have plotted each case study. Below, I discuss and justify the position of each of the case studies.

Figure 5.1 A theoretical framework for the use of Levelled Assessment Tasks



Originally, when making sense of the data, I had two continua; one for teaching and one for learning. The continua had summative and formative characteristics at each extreme, with the 'typical use' in the middle. After several attempts of representing the data along these and further analysis, I realised that the teachers' values towards the LATs may influence the pupils' values towards the tasks, so the use of a graphical representation would be more appropriate. Since, I believe that the teacher influences the pupils; the teacher's approach is placed along the x axis and the pupils on the y axis, following mathematical conventions of graph construction.

When coding the data the themes of values of the teachers and pupils towards the use of the LATs emerged from the data. This was whether they perceived the LATs as a summative activity or a formative activity.

5.1 Summative case study

Anecdotally, I have talked to teachers that use the LATs in test conditions. Pupils are expected to complete the LAT in silence and the teacher marks the resulting work, assigning a level. I was unable to observe such a lesson, but it is fair to assume that no learning takes place as the task is used as a testing exercise of what pupils know at that time.

The Summative Case exemplifies some of the characteristics of when the LATs are being used as a form of replacement summative test. The main themes that emerged were that the teacher did not have an overall understanding of levelness nor arguably the levels themselves, the pupils were aware of levels being used as an outcome, some limited learning did take place through doing the task because the pupils could talk to one another.

The teacher in this situation saw the LATs as an exam as illustrated by this part of the interview:

77	I	What is the purpose of doing the task with them?
78	T5	I'd say, from reading through it, it is their ability
79	T5	to interpret a structured question. Go through it to follow
80	T5	the information in it that they understand into a
81	T5	diagrammatical form.
82	I	What about the general use of these tasks?
83	T5	It enables pupils, when they get more
84	T5	comfortable with them, or when more able to self reflect
85	T5	to look specifically at the information they need to
86	T5	provide. In that respect I think it becomes easier for
87	T5	pupils once they understand that.
88	I	Is there anything you can do as a teacher to aid
89	I	this?
90	T5	Yes. Certainly tomorrow, I can run through: this
91	T5	is what it is asking for and these are the keywords. And
92	T5	that is primarily it, it is important for them to get the
93	T5	hang of that. Once they get the hang of it, it is easier for
94	T5	them to get the marks, but until they get the hang of it,
95	T5	they can put a lot of effort in and miss the marks… I
96	T5	would say that this is an issue with this. But that is how
97	T5	exams work! What can you do?

This conversation demonstrates that the teacher perceives the tasks as an exam, tasks that pupils have to learn to do so that they can 'get marks.' The statement about pupils 'need to self reflect' (line 84) at first could hint towards the idea of self assessment, but it appears that the teacher is very focused on the outcome of the task (getting better marks) and not the formative process of improvement and pupils developing an awareness of *how* to improve and having the opportunity to make that improvement.

In this case study it has been found that the teacher who uses the LATs summatively seems to focus on the outcome via marks, treating the LAT as a test (Conclusion 1).

Pupils in this case study were able to discuss levels, but limited to outcomes on the Level Ladder. There was no overall concept of what it meant to get particular levels. There was no opportunity for pupils to improve, nor did the teacher discuss levels or refer to the Level Ladder during the lesson. The levels were very much an outcome of the lesson. The pupils discussed the following when I asked them about levels:

121	Interviewer:	What [level] do you think you have got?
122	Imogen:	I think I may have got a 3 because I haven't
123		really done that much, I have just drawn the front of the
124		postcard really. So maybe a 3
125	Interviewer:	Ok. You have got an arrow there with a key
126		word on it.
127	Jessica:	Level 3 as well.
128	Kyle:	Well I think I have got, most probably a 3 but
129		maybe a 4.
130	Interviewer:	Which bits do you think will be maybe a 4?
131	Kyle:	Well when you draw these, it is definitely a 3
132		and gravity is like balanced and you get a 4 if you know
133		they are balanced.
134	Interviewer:	Well done, you did that off the top of your head
135		as well.
136	Liam:	I think a level 3, maybe just on the outskirts of a
137		4 because I didn't put much information really. I should
138		have put why the arrows are going up or down.
139	Interviewer:	Have you all used a Level Ladder before?
140	Note	[All agree that they have]

It appears from this that they were aware of levels and what they were aiming for, but not what the levels meant. Liam (lines 136-138) thought he had level 4 because he had not included ‘much information’ or not ‘put why the arrows are going up or down.’ The former is very generic and not level specific, I would expect a child who had an understanding of levels to talk about ‘using scientific keywords at level 4’ or using the idea of ‘balanced forces’ at level 5.

In this case study it has been found that where the teacher uses the tasks summatively, pupils associate levels as 'marks' and do not appear aware of the type of skills required to get a particular level (Conclusion 2).

5.2 Formative case study

If I now consider how a task can be used at the opposite extreme, where both the teacher and pupils value the formative processes when using the LATs, I am able to contrast the two case studies. This case relates to the 'formative approach' described by my 'pen portrait' (Table 4.2).

The Formative Case is an example of where teacher has fostered a culture that is focussed on the processes involved with learning science through the NCLDs. Many of the pupils in the class demonstrated that they understood what was required to achieve level 5 in a general sense and also understood that the focus of the lesson was on making an improvement, not demonstrating what they know.

In this case, the teacher used modelling as a way of conveying expectations at each level. She had prepared slides for the Interactive White Board (IWB) that showed the expectations at each level and was able to share the characteristics of each level. As she recognises, the pupils are working towards the concept of 'levelness'.

244	I	Do they have the idea of what a level 5 is?
245	T1	I think they are getting there, like me, I think
246	T1	they are a couple of steps behind…Maybe not
247	T1	what a level 5, maybe who is a level ladder, which
248	T1	student is a level 5.
249	I	Do you think that they are aspiring to move up
250	I	the ladder?
251	T1	Yeah and even the ones who say 'ooh, I don't
252	T1	care', I think that they see that they can move up
253	T1	with this sort of activity and you have to obviously
254	T1	to establish the ethos that it is good to learn and all

255	T1	the rest of it…no I do think they aspire to move
256	T1	up, but not be uncool at the same time.

The teacher recognises that a lot of the preparation of a lesson when using a LAT is based in engaging with the Level Ladder.

284	T1	I think you just have got to do the preparation
285	T1	beforehand you know, it's making those flip charts
286	T1	that made me fully understand the task, therefore
287	T1	in the lesson I am really comfortable about giving
288	T1	the advice whereas if you have thought 'brilliant
289	T1	It's one of those worksheet lessons – I don't have
290	T1	to really plan it-its already there for me' and you
291	T1	give it out, actually you don't really know what
292	T1	you are looking for in different levels.

This means that the teacher is ready to move around the classroom, identifying the key aspects of each level that pupils have achieved and ready to help them to consider their 'next steps.' Armed with this as a framework of conceptual development, the teacher is able to target their teaching towards individuals.

A striking observation from the pupil interview was that all the pupils interviewed identified a concept that was related to the progression in the Level Ladder as their 'best bits.'

10	Mac:	I think my best bit the particle diagram because I
11		find it really hard to write so drawing stuff is more
12		easier.
13	Interviewer:	Right…ok...thank you
14	Isobel:	I think my best bit was the saucepan...like...diagram
15		thing coz I done the heat arrows and thing going up the
16		spoon and stuff .
17	Interviewer:	Brilliant…ok...thank you
18	Teagan	I think my best bit was the saucepan because it's
19		labelled and it shows what different materials they all
20		are.

Mac identified the particle diagram, Isobel the ‘heat arrows’ and Teagan the materials (conductors and insulators). These are all linked directly to the descriptors in the Level Ladder for that task (Figure 4.2). This is in stark contrast to pupils in other case studies who generally focussed on their presentation (see the Transitional and Typical Cases).

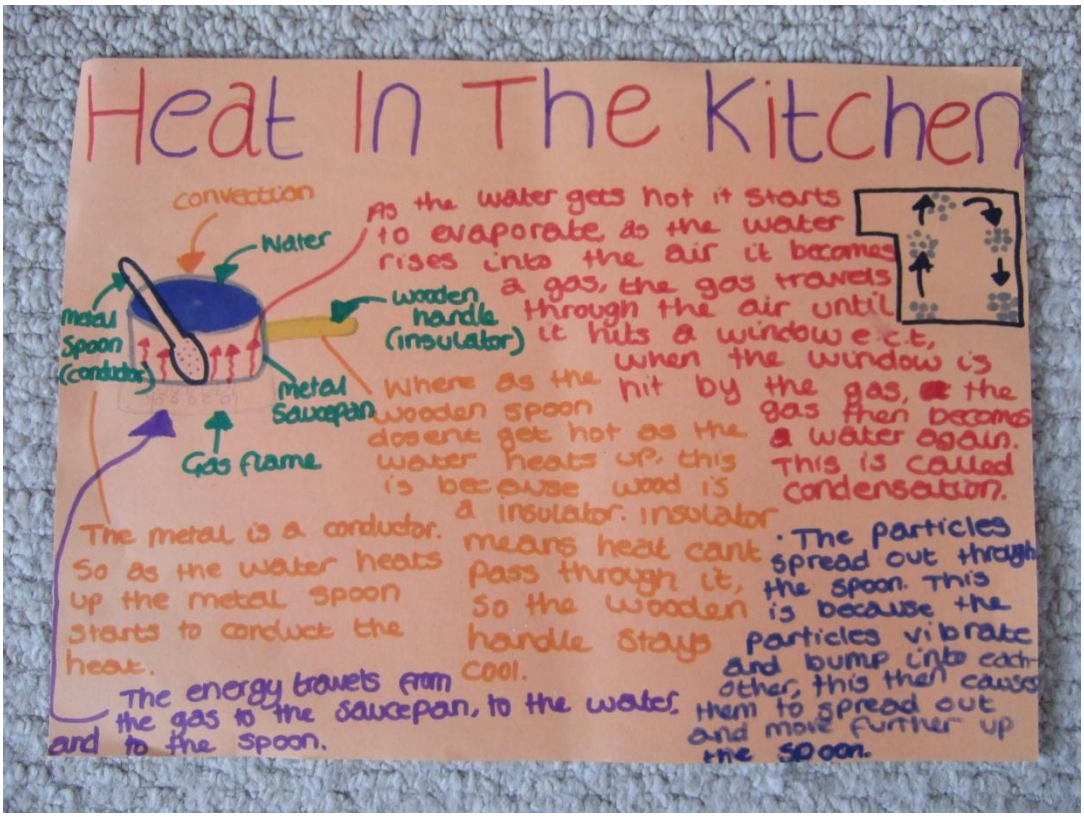
In this case study it was found that the teacher and the pupils had an overall view of the requirements for each level (Conclusion 3).

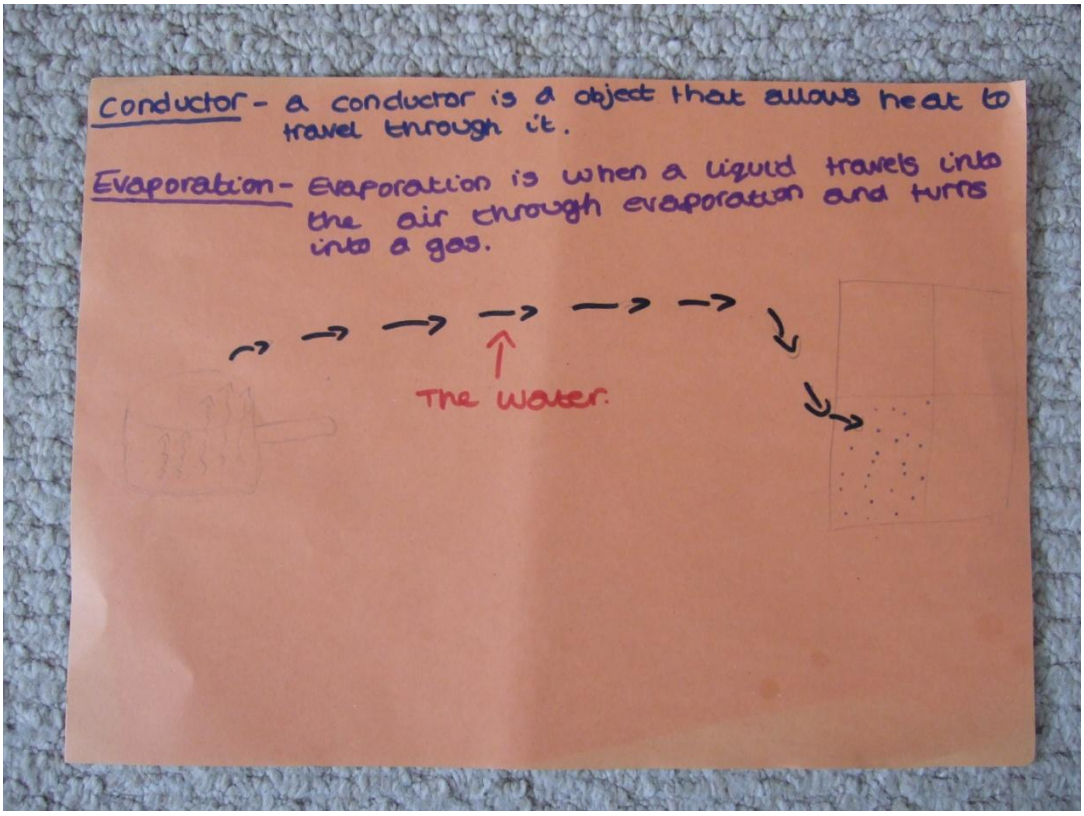
The key theme that emerged from this case was that pupils talked about their challenges and improvements based on the conceptual skills set out by the Level Ladder.

56	Isobel:	Um I found it difficult to explain um what how
57		the heat travelled like through each bit because I didn't
58		know whether to use particles. I am good at diagrams
59		but to explain it in words I found that difficult to be able
60		to do.
61	Interviewer:	Right ok and what did you do to help you do it better.
62	Isobel:	Well I like looked at the example when miss like
63		showed them but after that I'd already finished anyway
64		but I could see how I could improve.
65	Interviewer:	Ok...did you take the opportunity to improve or did
66		you just leave it as you had done it before?
67	Isobel:	I left it as I had done it before because it was the end
68		of the lesson.

It appears from Isobel, that the pupil used the visual models that the teacher presented to guide their progression, rather than the Level Ladder *per se*. In this class 92% of pupils claimed to use the Level Ladder to help (Table 4.8.Q2), so she was not necessarily typical. However, compared to the adjusted aggregates, 23% more of this class claimed to use the Level Ladder to help them. In this case study, slightly more pupils stated that they used the Level Ladder to improve, compared to the pooled results from the pupil questionnaire, 82% compared to 63% in the aggregate (Table 4.8.Q4).

Figure 5.2 Isobel's Work





83	Interviewer:	What bits did you improve in the lesson?
84	Teagan:	I improved this diagram because when I saw it on
85		the different levels on the whiteboard, I knew I that
86		would've had to explained how it travelled, so I drawed (sic)
87		how it evaporated from the saucepan and turn into gas
88		and then if it hit the window it turned back into water
89		and this is like condensation.
90	Interviewer:	So what does this diagram show?
91	Teagan:	That shows the gas travelling through but that never
92		got finished….and then the water particles turn back
93		from the gas to the water.

With using the modelling approach, I had concerns that pupils would just be copying their teacher's model. If this was the case, I questioned to what extent did that support their learning. However, at least in the following example, the pupil had only added her improvement when she understood it (Figure 5.2 Isobel's Work):

95	Isobel:	My improvement, well I did a couple of things
96		because when I looked at the board I only got a level
97		5/6 so to improve I done…what they called?…heat transfer
98		diagrams and I also done a diagram of the water
99		particles changing from the liquid to the gas to show
100		evaporation and I added in some…like...coz I had just
101		written down the words and the meanings of the words I
102		had to link them all together to be able to show
103		This...how heat travels.
104	Interviewer:	Now that diagram you have got there of the
105		convection is very similar to the one on the board, did
106		you copy that?
107	Isobel:	Yeah I added it in after I saw it on the board.
108	Interviewer:	And do you know what it means, are you able to
109		explain it?
110	Isobel:	Kind of.
111	Interviewer:	Kind of. Right. Coz I saw a few people had done that
112		and put it down and I wondered how much they actually
113		understood it. Would you be able to tell me anything
114		about that diagram?
115	Isobel:	Well that's the water particles just like as they
116		are starting to heat up and then they spread out coz they
117		are vibrating and as they spread out they get lighter and
118		lighter and then as it cools back down again it goes back
119		to the bottom.

120	Mac:	The diagram coz um I was looking at wood and
121		everything and I noticed that I hadn't done a proper
122		particle diagram showing how heat goes through all the
123		particles and all that. And then Miss showed me how to do it on
124		the board so I just took it off the board really.
125	Interviewer:	And what does your particle diagram show there then?
126	Mac:	There's a heat arrow and they are like all vibrating
127		and showing how heat is passed along.
128	Interviewer:	And how is the heat passed along in that situation.
129	Mac:	That they all hit each other and the heat goes along.

In this case, where the pupils and teacher have formative values, the pupils focussed on their improvement (Conclusion 4).

In this case only a third of the class strongly agreed or agreed that another pupil helped them when they got stuck. A further third ticked 'unsure' for this question (Table 4.6.Q7). When I asked the interview group of pupils whether they helped each other with the task, they said:

139	Isobel:	Yes. Because you can kind of tell them what they
140		need to improve…and you can discuss what you have
141		already got and what they have got and you can like
142		swap ideas.
143	Interviewer:	Can you give me any examples of anything you did
144		today that you got from another student in the room?
145	Teagan	That purple writing about the energy travels, I was
146		asking Emily how the energy travels up and down the
147		spoon and reaches the top?
148	Interviewer:	Did you find you learnt from talking to Emily?
149	Teagan	Because Sarah sits just over there, we like discuss things
150		all the time, just about what was that last lesson that we
151		learnt, and what are we meant to be doing for this and
152		what does this mean and stuff.
153	Mac:	I talked to Tim who is almost right next me. He said
154		that you had to mention that that bit was wooden and all
155		that.

It is probable that some pupils in this group did gain some information about how to complete the task from each other, above three explicit examples are given. Why were so many uncertain if they had been helped by another pupil?

156	Interviewer:	How much did you use the Level Ladder to help you?
157	Teagan	The most I think, I keep flicking back to it to see if I
158		have missed anything out.
159	Mac:	I just look at it at the beginning and then go ok I am
160		doing that and just carry on.
161	Interviewer:	Did you look it again towards the end at all. So you
162		used it a lot A andB, but you C didn't. So how do you
163		know to improve, what did you rely on?
164	Mac:	The board, because we normally do that every time
165		we have done one of these she normally gets out how to
166		improve it and all that, what levels you have to do and
167		all that.

This conversation adds to the pupil questionnaire and my observations:

In the formative approach case, some pupils supported each other's learning through engagement with the Level Ladder (Conclusion 5).

5.3 Typical case study

This case exemplifies what I now see as the ‘typical’ use of the LATs (Table 4.2). The online questionnaire exposed that on the whole LATs were not used as they were intended, notably that they were used as a replacement task and that pupils were not given the opportunity to make improvements.

An apparent theme that developed throughout the observations of the tasks and pupil interviews in use was that the pupils had the view that presentation was more important than the content of the task. This manifested itself in two ways: over emphasis on drawing a title or relating achievement to the neatness or tidiness of their work.

While observing the lessons, I noted that in four of the five cases there were at least a few pupils that spent most of their time drawing an elaborate title for the task at the expense of any attempt of the task. However, these were to a different extent in each case.

In the Typical Case, although John recognised his ‘best bit’ as the food chain, it was cited because it was ‘easy’ and that they had not ‘scribbled out or it’s tidy’.

45	John:	I stuck it on my diagram of what eats what.
46	Interviewer:	Okay.
47	John:	I find that the best because it's easy and, like,
48		really easy to understand, and simple.
49	Interviewer:	Okay, good. So what have you got there?
50		You've got four food chains there and you chose that
51		one. Why did you choose that one?
52	John:	Because it's the only one that I haven't
53		scribbled out or it's tidy.

There was no value on achievement, something that they had learnt or personal improvement. In the same group, although Amanda acknowledges that the ‘information’ is more important than the title, she does not refer to the Level Ladder at

all and John identified the layout and ‘appeal’ of the work as the significant factor in deciding their ‘best bit’:

153	Interviewer:	Okay, that's the
154		next question, so if you were to start this lesson right
155		from the beginning again, what would you do
156		differently, to improve, I suppose, on what you've
157		done so far? And shall we start with Amanda this time?
158	Amanda:	Well, if I started again I would probably start
159		researching, like, write all the writing first, and then
160		I'd do the food chain and then the title, to see how
161		much room I'd got left.
162	Interviewer:	Okay. Why would you do the title last, if you
163		did it again?
164	Amanda:	Because if I did, like, the title first, it might use
165		up all the space and I wouldn't have enough room to
166		do all the information.
167	Interviewer:	Okay. And is the information more important
168		than the title?
169	Amanda:	Yeah.
170	Interviewer:	Yeah? Why, why is that?
171	Amanda:	Because you can find out more and it's more
172		interesting than just looking at the title.
173	Interviewer:	Right, okay. Fair enough. And what about
174		you, John? What would you do if you were to do it
175		all over again?
176	John:	I think I might have sort of presented it a bit
177		neater.
178	Interviewer:	Okay, right. But you were happy with what
179		you handed in?
180	John:	Yeah, mm.
181	Interviewer:	Would you get marked down for not handing it
182		in? For handing it any neater? Would you get any
183		better marks?
184	John:	Yeah, because it's a bit all over the place, in a
185		way.
186	Interviewer:	Okay, so you're talking about the way it's laid
187		out?
188	John:	Yeah.
189	Interviewer:	Okay, so…
190	John:	I'd lay it out, yeah.
191	Interviewer:	Okay. So what would you… could you give
192		me any examples there?

193	John:	Well, I'd probably do my food chains a bit
194		more together and neater.
195	Interviewer:	Okay.
196	John:	Apart from this one. And my lists at the top,
197		maybe I could have done it in [inaudible 08:44].
198	Interviewer:	You've done them in little columns though,
199		haven't you?
200	John:	Yeah.
201	Interviewer:	Yeah, that's fair enough. Okay, thank you.
202		And Mark, what would you do differently?
203	Mark:	I would have read through the sheet a bit, like,
204		clearer, because I think I missed out a section, a
205		small section.
206	Interviewer:	Right, okay.
207	Mark:	And then, so… but if I'd missed it out before I
208		would have redone it underneath, but obviously I
209		didn't have enough time, as I just spotted that I'd
210		missed that bit just at the end of the lesson.

I explored with John how this was related to levels. When talking about levels, Mark shows that he has not really engaged with what a ‘level 5’ or a ‘level 6’ means, despite having the task sheet in front of him with the Level Ladder available to consult. This raises questions about whether pupils in this class really understand the requirements of the task and the differences with the teacher in the Summative Case, where pupils were developing a concept of levelness.

213	Interviewer:	Do you know what level you think you might
214		have for your work today and why? And if you don't
215		know, that's fine as well, okay, but if you just pull a
216		level out the top of your heads, what have you got
217		and why, from what you've done today? So, we'll
218		start with John.
219	John:	I think I might have got a 4.
220	Interviewer:	Right.
221	John:	'Cause my presentation isn't good, but I've got
222		basically most of it down.
223	Interviewer:	Mm.
224	John:	I think I'm probably going to get about 4.
225	Amanda:	I think… I've no idea, but…

226	Interviewer:	No idea?
227	Amanda:	Probably a 4 or something.
228	Interviewer:	Okay. How would you know that you said a
229		four or something? What's that based on?
230	Amanda:	Like, the bold title and the information that's
231		on it.
232	Interviewer:	Okay, right. And the information on the side's
233		about what there?
234	Amanda:	Food chains, like a big food chain.
235	Interviewer:	It's a big food chain.
236	Amanda:	At the side.
237	Interviewer:	Okay, excellent. Fair enough. And Mark?
238	Mark:	I would sort of give myself level 5 but only just,
239		because it could have been better but there's a lot of
240		information on there and I did miss out that section.
241	Interviewer:	Okay.
242	Mark:	So I guess I would lose a file mark for that,
243		but…
244	Interviewer:	Okay, fair enough. But what makes you think
245		that would be level 5 then?
246	Mark:	Because at level 6 it would be a lot more, like,
247		sophisticated than this.
248	Interviewer:	Right.
249	Mark:	And, well, what I've done here, 'cause I didn't
250		really understand to begin with, is very, like, basic,
251		but yeah, it has a lot of information in it.

The teacher in Typical Case recognised that some pupils did spend too much time on the title and attributed it to the pupils not being clear about what they were doing.

23	T:	So probably if I had given that input [a cover teacher took the previous lesson], they
24		would have got going on it faster, but still, despite the
25		fact that you've explained it clearly, some kids might
26		not listen clearly, or when they actually look at the
27		task again, they kind of… oh, I don't get it, what do I
28		have to do? But a lot of them, as soon as you just
29		pointed them in the right direction, got going. Some
30		just got straight into it. It's lovely really. I think it's
31		the amount that they have to get into their work to get
32		the right level. A lot of them spent nearly the whole
33		lesson just drawing a food chain of a couple of

34		organisms. Maybe they were distracted or they
35		wanted to do it really carefully in colour. They
36		seemed to spend loads of time on the presentation
37		and not enough actually getting information down.
38		So I will have to spend more time on this, which I
39		hadn't planned to do.

Although the teacher puts the blame on the previous lesson, I would suggest that a shift in focus during the lesson would avoid this. Firstly, the teacher could simply state that pupils should only write a short title and concentrate on working through the Level Ladder. Secondly, the teacher could encourage pupils to try to seek help by consulting their notes, a text book or each other. Finally, the teacher could emphasise to the pupils that it is important that they learn something during the lesson and place value on that improvement, not just on the level that they get.

From the pupil questionnaire, 74% of pupils agreed or strongly agreed that they used the Level Ladder to help them which is typical compared to the adjusted aggregate of 73% (Table 4.8 Q2). This group did not appear to value working with each other on the task, with 56% disagreeing that they another pupil helped them when they got stuck, 13% lower than the aggregate of 43% (Figure 4.8.Q7).

In this case, when the task is used in a 'typical' approach pupils often value presentation over learning (Conclusion 6).

The teacher has the intention to encourage pupils to self assess and identify improvements during the following lesson:

97	T:	Well, they haven't finished them, so next time I
98		see them, I'll give them back out and ask them to – if
99		they've finished it, I'll ask them to try and level it, and
100		then suggest improvements, and then put those
101		improvements in to get to the next level. If they

102	haven't finished it, I will give them time to finish it, so
103	we do about half a lesson on it.

However, the pupils did not seem to engage in this very well when I interviewed them (see above). The teacher treats this as an 'add-on', rather than an integral part of the session. Contrast this with the Formative Case where the teacher has the lesson focussed on the improvement process, rather than the outcome of a level. This, I feel illustrates that the teacher has 'formative intentions' but is not fully embracing the concept of formative assessment into his pedagogical approach.

When using LATs, typically teachers focus on the outcome level, rather than the process of gaining that level (Conclusion 7).

5.4 Transitional case study

As I have suggested, some simple changes in approach by the teacher in the Typical Case could have an impact on student values. This may be illustrated by a comparison with the Transitional Case. Both lessons had a similar structure, neither giving explicit emphasis on improvement. However, the teacher in the Transitional Case used another strategy to enhance pupil engagement with the Level Ladder by circulating during the lesson and actually ticking pupils' Level Ladders.

In the Transitional Case, the pupils did offer knowledge gains as their 'best parts' of their work:

40	Brian:	Mine was talking about what happened at the
41		beginning, and what happened at the end, does it,
42		like, burn with the oxygen and magnesium.
43	Interviewer:	Good, okay. And why is that? Why do you
44		think that is the best bit?
45	Brian:	Because it looked like the best part of that,
46		and I liked talking about it, as well.

47	Interviewer:	Right, okay, so you like talking about it as well.
48		Brilliant, okay, thank you. Cathy?
49	Cathy:	I understood the best, the most, and it was
50		easy, most, and I got a higher level.
51	Interviewer:	Okay. So which bit was that? Sorry, that's
52		the… the bit underneath there.
53	Cathy:	Yeah, Level 5.
54	Interviewer:	The bit that you got Level 5 for. Okay.
55		Including the diagram there?
56	Cathy:	Mm.
57	Interviewer:	Yep. So that was the best bit in your opinion?
58		Okay, good, thank you. Dawn?
59	Dawn:	My one was mostly all of it really because I
60		liked the way, when he showed it to us, I understood
61		what he [the teacher] did, and I enjoyed describing how he did it,
62		and we had to look at the sheet to help us, and that
63		helped me a lot.

The pupils are talking about the knowledge gains in science and appear to really enjoy describing and explaining what they understand. In contrast to the Typical Case, where 84% of pupils claimed to have learnt something new, only 50% claimed this in the Transitional Case (Table 4.8 Q5). There was far more uncertainty amongst pupils in the Transitional Case as well, as 28% unsure if they had learnt anything new, compared with the Typical Case where only 4% were uncertain.

The reason for this different emphasis maybe due to the fact the teacher circulated, assigning a level to different parts of pupils' work during the lesson, as he describes:

25	T:	.... I go round and I scribble
26		numbers – you know, I try and do that, and
27		that really motivates them, because they…
28		I've got a 3, you've got a 4… or, you've got
29		a 4 plus. And so that I find really very
30		useful.

So, the pupils could connect very clearly between the teacher's values of levels and the Level Ladder. From interviewing them, pupils in the Transitional Case were able to talk about levels, and voluntarily referred to the Level Ladder during the interview.

198	Cathy:	I had to use Level 6 target.
199	Interviewer:	Number two there, yeah.
200	Cathy:	So it's going why the substances can be
201		classified as in element, compound or a mixture.
202		And I chose that one because I wasn't quite sure, so
203		I wanted to challenge myself to try… to make it a
204		target, so when I get that, I can get another harder
205		target.

Paradoxically, in this class had a high 39% of pupils claiming that they did not use the Level Ladder to help them, compared to 4% in the aggregate. In addition, 44% of the class disagreed that they used the Level Ladder instead of asking their teacher, compared to 25% in the aggregate (Table 4.8 Q6). Maybe because the teacher was talking with the pupils throughout the lesson, the pupils did not need to or perceive that the Level Ladder supported them, but instead their teacher.

From my observation, it was clear that the teacher had developed a very good relationship with the class and a majority of them enjoyed his lessons. Again, in contradiction, this case scored lowest in the questionnaire, when pupils were asked if they enjoyed doing the task.

I have positioned this case study on Figure 5.2 tentatively in line with the Typical Case, but with the learning slightly more formative. From my observation, I was certain that this case had the lesson structure and the teacher's approach met my 'typical approach' criteria in Table 4.2. However, I became convinced from the pupil questionnaire and interview, that there was a pupil focus on learning conceptual knowledge but uncertainty about the extent to which the Level Ladder was used to support this. Pupils

in the interview readily used the Level Ladder when answering questions about learning, levels and improvement. I wonder if they are at the point of taking the Level Ladder for granted and are unaware of how much they use it. This is why I have called this the 'transitional' case; it appears to be moving towards a formative approach.

In this case, the teacher's emphasis on the use of the Level Ladder influenced pupils' values on gaining knowledge and understanding (Conclusion 8).

5.5 The Mismatch case study

The teacher in the Mismatch case adapted a LAT to produce a lesson where the pupils worked collaboratively to ascertain if a patient's lifestyle was healthy and what recommendations they would make to improve the patient's lifestyle. This was a top set, very articulate pupils and obviously all successful learners.

The pupils really enjoyed doing the task, with the highest enjoyment score in the pupil questionnaires (Table 4.6 Q3) of 92% (aggregated 59%) and 76% claimed to have learnt something new, compared with the aggregate of 78% (Table 4.8 Q5).

The teacher used the task as an introduction to the topic.

48	T	... And it fitted in perfectly with my
49		scheme of things. I wouldn't use them just to prove a
50		point. It's gotta be because I was going to do something or
51		we have done something. I think in a way they would
52		work better as a plenary session for a lot of topic areas
		...
59	T	... They have got the idea, they know the words but you
60		don't know what emphysema is or what alveoli are, or
61		whatever it is. I can go through that and for me that means
62		that I know that need not do the lower end of the
63		topic….So it will be a teaching tool and a feed-forward
64		activity. Something for me.

The most startling aspect of this case was the mismatch in the teacher's ideology for formative assessment and the pupil's expectation of what teaching and learning is. Something that the teacher noted herself:

3	T	I would say it probably doesn't in the fact that is
4		not the way they like to learn. So although I can use the
5		Level Ladders to help me establish that I'm hitting the
6		right target levels at the top end. For them, they find them
7		a bit too open ended. They like black and white answers,
8		they like to have the mark scheme, their ideals Assessment
9		for Learning tool would be have a test, there's the mark
10		scheme, it has been marked by the teacher so it has a level
11		that they trust. Then they look at the mark scheme and say
12		'oh yeah, I see where I have gone wrong and I know
13		where to go now, I won't do that again.' They really like
14		that, as a whole group they like that, almost without
15		exception, that would be their choice if you gave them any
16		assessment tool.

The pupils were not clear about the function of the LAT itself. In a way they were treating it as an examination. I asked them how they had selected the work that would be presented to the rest of the group:

273	Interviewer:	So you did have some idea of the level of that
274		work going up.
275	Rose:	No. We just thought it had more keywords that
276		would boost our level.
277	Toby:	I find this keyword idea quite iffy if you ask me. I
278		mean it's like saying, 'Use these words that proves you
279		know everything.' You are just writing out some words
280		they have told you to write out. What does that prove? Do
281		you know what these words mean? It hasn't proved any
282		initiative.
283	Rose:	If they were going to give keywords, then I could
284		give a description of each one so you could write it in the
285		right text.
286	Toby:	It is saying using these keywords makes you are

287		automatically a very clever person.
288	Interviewer:	Using a keyword is only a level 4 skill.
289	Toby:	I would say , instead of use these words in your
290		explanation, you could say…give... a more broad ‘consider
291		these topics’.
292	Interviewer:	But it does in this one, say use the keywords.
293	Rose:	Coz it says on them to use the keywords, but if you
294		don’t know what some of these mean, and if this is a test
295		for what we did know. If you don’t know what some of
296		these mean, it makes you seem like you should know. By
297		putting them in there puts a lot more pressure on you…
298	Toby:	If I use a better word, do I get a worse mark?
299	Interviewer:	Well you tell me.
300	Toby:	Well if they are there must be a reason for them or
301		are they just there to fill up space?
302	Interviewer:	[laughs] Or to prompt you maybe? So you would
303		prefer no keywords?
304	Toby:	If you would like to give the students some ideas,
305		you could say these are some of the things you can think
306		about. Not, USE these key words.
307	Rose:	Coz we have used these sheets before in previous
308		years in lessons, it will say use the keywords to boost your
309		levels. And so if you don’t know what some of the
310		keywords mean, you feel really like…
311	Interviewer:	But what can you do about that?
312	Rose:	If we are doing it in a test conditions we wouldn’t
313		have a book there to look it up. But as we have it here, we
314		could look through all the books and as we didn’t have
315		enough time we couldn’t look through all the books and
316		we might not find it. And as the books we were looking at
317		were from the GP and they are not going to have all the
318		information about the alveoli. It doesn’t go deep into
319		things. So if you are going to do it, use text books, if they
320		expect us to use keywords, we should be given text-books
321		to help us.

The pupils were genuinely confused about the Level Ladder and this type of task. This class were a high achieving group and displayed the characteristics of confident, intelligent pupils. Although they were familiar with the tasks, it is clear that they want

to understand the 'rules of assessment.' They are clearly focussed on wanting to achieve a high level, in this case a 'levels' focus rather than a 'levelness' focus. Toby talks about what it means to be clever and Rose thought that using more key words would secure a better grade. I found it interesting that Rose talked about test conditions versus open book approaches, the only pupil in all the interviews that did. It seems to me that she values recall over understanding, perhaps another aspect of the levels-levelness dichotomy.

However the collaborative aspect was beneficial to these pupils, with 60% claiming that another pupil helped them, compared with the aggregate of 42% (Table 4.8 Q7). So although they recognised the value of the experience, it appears to me that they challenged the teacher's formative approach since they were usually successful in topic tests and summative assessment approaches.

In this case study, high achieving pupils challenged the formative approach of the teacher compared to the summative approach they were used to. (Conclusion 9)

5.6 Summary of conclusions from case studies.

From the data collected from the five case studies, I have been able to make nine conclusions from individual cases, using Figure 5.1, the graphical representation of a proposed relationship between teaching approach and learning. Some of these claims have been developed by doing some comparative analysis between the cases.

From the Case Studies the conclusions are:

1. The teacher who uses the LATs summatively seems to focus on the outcome via marks, treating the LAT as a test
2. Where the teacher uses the tasks summatively, pupils associate levels as 'marks' and do not appear aware of the type of skills required to get a particular level.
3. The teacher and the pupils had an overall view of the requirements for each level.

4. Where the pupils and teacher have formative values, the pupils focused on their improvement.
5. In the formative approach case, some pupils supported each other's learning through engagement with the Level Ladder.
6. When the task is used in a 'typical' approach pupils often value presentation over learning.
7. When using LATs in the typical approach, teachers focus on the outcome level, rather than the process of gaining that level.
8. The teacher's emphasis on the use of the Level Ladder influenced pupils' values on improving knowledge and understanding.
9. High achieving pupils challenged the formative approach of the teacher compared to the summative approach they were used to.

From these conclusions I want to make some generalisations about LAT use in order to address my original research questions. How are the tasks used and to what extent they support teaching and learning?

Bassey's (1999) educational case studies approach through fuzzy propositions has provided me with a framework within which to make tentative claims to knowledge and the structure in which to propose generalisations based on the research I have carried out. The claims are importantly language dependent in the way in which I make my claims to knowledge. My conclusions above, I defend as conclusions consistent with the evidence I have observed and collected. These can be presented with a high degree of certainty. So I was able to use the language:

In this case ... it has been found that...

Through using fuzzy generalisations, Bassey (1999) contends that the language for making generalisations should be more tentative, for example:

In some cases it may be found that...

Interrogating my nine conclusions and the online data has led me to present four fuzzy propositions about the use of LATs in the science lesson.

Fuzzy proposition 1:

In some cases it may be found that teachers who focus on ‘levels’ use the LATs as a type of exam and the Level Ladder as a mark scheme, valuing the marks or scores achieved rather than using the Level Ladder as a scaffold for learning.

Fuzzy proposition 2:

In some cases it may be found that teachers who value levels as outcomes cause pupils to value presentation and final outcome rather than improvement to knowledge and understanding.

Fuzzy proposition 3:

The language that the teacher (and pupil) uses when using the LATs is essential in creating a ‘formative’ learning culture.

Fuzzy proposition 4:

Classroom climate of ‘level-focussed’ teachers is concerned with outcomes as opposed to the learning process of improvement.

The observable elements of these propositions are displayed in Table 5.1, derived from the analysis of the Case Studies.

5.7 Fuzzy Generalisations

Considering the evidence gathered within the framework proposed in Figure 4.1, it is possible to settle upon three fuzzy generalisations. I consider this framework to be a starting point upon which any similar case study could be placed. If another five case studies were to be carried out, I would expect that the fuzzy propositions (above) may be further refined and the fuzzy generalisations would be further supported. In the spirit of grounded theory, it is unlikely that the data is saturated at this point, but more case studies would add to this.

Fuzzy Generalisation 1:

Teacher attitudes to the LATs may influence pupil attitudes to the LATs.

Fuzzy Generalisation 2:

Teachers with a ‘big picture of levels’ may be more likely to use LATs formatively.

Fuzzy Generalisation 3:

Teachers who engage pupils with the notion of levelness may be more likely to improve conceptual development of pupils.

Table 5.1 Characteristics of classrooms using a ‘Levels’ or ‘Levelness’ approach

Levels (Outcome focus) Climate	**Levelness (Process focus) Climate**
LAT used as a type of test or exam, feedback is limited.	LAT is used as an assessment tool to recognise current understanding and support further learning.
Levels used as ‘marks’ to be recorded.	Levels used as a language to communicate expected outcomes.
Use Level Ladder as a mark scheme.	Use Level Ladder as a scaffold for cognitive development.
Pupils more likely to be focused on presentation rather than conceptual gain.	Pupils more concerned on conceptual gain, where they are now, how to improve.
Values the outcome of the task.	Values the process of doing the task.
Teachers may place emphasis on pupils’ effort, presentation and marks.	Teachers may place emphasis on pupils’ current knowledge and improvements.
Teachers use the criteria as a series of unrelated points to achieve.	Teachers and pupils have an understanding of the meaning of each level and the relationship between them.
Language used by teachers and pupils is focused on achievement, getting marks, levels as marks	Language focused on what levels mean, assessing current understanding and finding ways to improve.

Next, I summarise the evidence that supports each of these three fuzzy generalisations. Fuzzy Generalisation 1 states that ‘Teacher attitudes to the LATs may influence pupil attitudes to the LATs.’ The online questionnaire provided insight to statistically generalisable attitudes of science teachers towards the LATs, which I was able to summarise in Table 4.1 as three approaches. It was the Case Studies that fleshed out these claims. The most apparent contrast was between the Summative Case and the Formative Case (Figure 5.1). The Summative Case showed that the teacher regarded the

LAT as a summative test, through both the lesson observation and the teacher interview. However, in the Formative Case the teacher regarded the LAT as an activity that could diagnose the level at which the pupils were working at and then assist the pupils with making improvements. In the Formative Case the pupils were able to talk about their improvements and levels that were similar to that of the teacher. Their attention was more focussed on improving their work compared to the pupils in Summative Case where the pupils regarded the levels and outcomes and were more focussed on presentation. In addition, the Mismatch Case added further support for this, but perhaps demonstrates that there is a 'lag' between teacher and pupil attitudes to the task. The teacher had far more 'formative' attitudes towards the task than the pupils, but in this case the pupils were challenging the nature of the LAT and compare it to their experiences of traditional tests (Mismatch Case, Pupil Interview). The main assumption with this fuzzy generalisation is that the teachers' values influence the values of the pupils to the LAT. This comes about by the fact that most of the case studies show some correlation between teaching and learning displayed in Figure 5.1.

For Fuzzy Generalisation 2 I claim that teachers with a 'big picture of levels' may be more likely to use LATs formatively. This was first suggested from the descriptive framework that I developed from the online questionnaire (Figure 4.1). The case studies added further evidence for this. It was clear that the teachers in the Formative and Mismatch Cases had a clear understanding of the generic expectations of each level and the teacher in the Summative Case did not. In the Transition Case, where the teacher went round the class assigning levels, had a good understanding of the levels to be able to do that however the tasks were not being used as formatively as they could have been. The pupils did not demonstrate a 'formative attitude' to the task compared with those in the Formative Case.

Finally, I determined Fuzzy Generalisation 3 as teachers who engage pupils with the notion of ‘levelness’ may be more likely to improve conceptual development of pupils. This was shown significantly in the Formative Case where pupils were encouraged to consider what each level would ‘look like.’ Its absence in the other case studies was notable. However, Transition Case pupils were able to engage with the Level Ladder and referred to it when talking about their understanding, whereas pupils in the Typical Case made little reference to the Level Ladder, but instead judged their work on presentation rather than conceptual improvement. In neither case have the pupils an awareness of ‘levelness’.

The claims that I can make from this research are:

- Generally teachers do not use the LATs in the way in which they intended.
- A majority of teachers do not use tasks as formatively as they could.
- From the case studies, the attitude and strategies of the teacher when using a LAT determine the values of the pupils towards the LAT.
- LATs do support learning in most cases.
- LATs can support teaching in a variety of ways, both summative and formative.

5.8 Replication invited

Bassey (1998) argues that case studies being singularities can add strength to any fuzzy generalisations or fuzzy propositions by replication. This is my invitation to others to carry out a similar case study of teachers using the LATs and analysing it to see whether it supports my theoretical framework or adds another descriptive or explanatory perspectives. I discuss the implications of this in Chapter 6.

5.9 Data audit

In the spirit of case study educational research and the invitation to others to replicate, I have endeavoured to store the raw data from this research in an accessible way, so that anyone interested could attempt to reanalyse and interpret the raw data. In addition the raw data displays my attempts at coding it, which means that when interrogating the research, a reader could scrutinise my coding decisions.

In a single spreadsheet, I have presented the raw data and where appropriate, codes.

The spreadsheet contains worksheets that display the following information:

Online Questionnaire:

- Case Study Lesson Observations (Cases 1-5):
- Case Study Pupil Questionnaire (Cases 1-5):
- Case Study Pupil Interview Transcripts and coding (Cases 1-5):
- Case Study Teacher Interview Transcripts and coding (Cases 1-5):

Throughout the thesis I have made reference to this data using a coding system, so that the source of data can be easily traced within the spreadsheet.

Bassey (1999) suggests that an audit certificate should be provided with the statement:

> *In terms of the evidence provided in this paper it is my professional judgement that the statement of empirical findings is based firmly on the data collected and that the enquiry has been conducted according to the ethical guidelines of respect for persons and respect for truth. (p.19)*

This is expected to be signed and dated by a supervisor or colleague. I feel in this case, the thesis is judged by two supervisors and two examiners who can make that decision based on my thesis and my defence of it during the viva process.

Chapter 6 Discussion

I have approached this thesis as a narrative of my development from being a classroom teacher to a researcher in education. Chapter 1 offers an account of what my knowledge and understanding of AfL in science education was at the start of my research. In keeping with a Grounded Theory approach, I relate my findings to the academic literature in this part of the thesis, not in a literature review in an early chapter. This reflects both GT and an honest portrayal of my research process.

The five case studies of teachers using LATs in lessons provided detailed insights into what was going on in these classrooms from my perspective, the teacher perspective and the pupil perspective. These case studies could be seen as representative of the Pen Portraits (Table 4.2) that I composed to describe the types of approach that teachers use when using the LATs. I characterised these three approaches as the teacher using the tasks 'summatively' where the teacher uses the LAT as a replacement summative test, 'typically', which is outcome focussed, but with some elements of formative approaches used and 'formatively' where the teacher is improvement focussed.

The Summative Case best represented the summative use of the task, but was not the most extreme version of this approach. Although I did not get to observe this, LATs are used in some schools as topic tests carried out in silence and in test conditions. The Transitional and Typical cases represented typical use of the tasks where the teacher had some formative ideals but both the teacher and pupils were focussed on outcomes. The formative approach was best represented by the Formative Case where both teacher and pupils were focussed on improvement and learning in the lesson. Finally, the Mismatch Case demonstrates a situation where the pupils, who are successful in a summative system, have an ideology at odds with their teacher's formative approach.

Many issues have been raised by this research and so in this chapter I have selected what I consider the most important and interesting issues to discuss. These issues are the extent to which the LATs are formative assessment tools, and the notions of 'levels' and 'levelness.' I then go on to explore three questions; should teachers communicate levels with pupils? Why do teachers use the LATs? and How can teachers be encouraged to use the LATs more formatively?

6.1 LATs and formative assessment

It is clear from my research that few teachers use the LATs in the way they were intended. If the online survey is a good indicator of the approaches used, it appears that a minority of teachers use the LATs for purely summative purposes, most use them in a pseudo-AfL approach and then a small number use them formatively – their intended use. This means that a majority of teachers do not use tasks as formatively as they could. The previous chapter illustrates that it is not the fact that teachers use the LATs that makes the lesson or learning experience formative, it is *how* the teacher uses the LATs.

As I explained in Chapter 1, when I developed the LATs my definition of formative assessment was based on a combination of Black and Wiliam (1998a) and the NSS interpretation (NSS, 2004). Now, in the light of my research, I reconsider what is meant by AfL and the relationship of the LATs to AfL. Black and Wiliam (1998a) define formative assessment as:

> encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged. *(ibid., p. 7)*

They recognise that there are many approaches that fulfil this definition. This is the intended purpose of the LATs; they provide a framework through the Level Ladder for

both teachers and pupils to guide improvement through feedback (Grevatt 2005a). The areas of practice that Black and Wiliam (1998b) identify as strategies that support formative assessment are sharing learning objectives, questioning, self assessment, peer assessment, written and verbal feedback and the formative use of summative tests. These were translated into practical approaches by the NSS. The LATs are intended to encourage the sharing of objectives and outcomes, opportunities for self and peer assessment and more effective feedback.

Bell and Cowie (2001b) carried out significant research of formative assessment in science education in New Zealand classrooms of 11-14 year olds. They defined formative assessment in this context as:

> the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning. (p.540)

Their definition, compared to Black and William's (1998a), does not contain the concept of 'feedback' *per se* instead focuses on triggers that can then lead to responses that can enhance learning.

Other research investigating formative assessment in primary schools was carried out at a similar time by Torrance and Pryor (1998), from which they identified two versions of formative assessment. These two versions they called 'convergent assessment' which aims to discover whether a learner knows, understands or can do a predetermined thing, compared to 'divergent assessment' which aims to discover what the learner knows, understands or can do. At first glance the LATs fulfil the convergent assessment criteria in that they require precise planning, tick lists and can do statements, an analysis of the interaction of the learner, the curriculum from the view point of the curriculum and the pupil is the recipient of the assessment. However, I would contend using the evidence from my research that some aspects of divergent formative assessment are achievable

through the use of the LATs, making them have divergent assessment characteristics. The tasks are open, they do yield insight to learners' understanding (e.g. misconceptions are highlighted), they are descriptive in their evaluation, but do require some judgemental and quantitative evaluation.

I have some issues with the theoretical implications of each type. Torrance and Pryor (1998) suggest that convergent assessment assumes a behaviourist view of learning, rather than a constructivist perspective. They also assume that it has the intention to teach the next step in a linear progression instead of in the zone of proximal development (ZPD) and see assessment as accomplished by the teacher, rather than a joint accomplishment of both teacher and pupil. I argue that science learning does have a general linear path for progression, especially in developing concepts. There have been a multitude of studies on conceptual development (e.g. Leach and Scott, 2002, Taber, 2006, Franco and Taber, 2009), the progression of learning models and modelling (e.g. Gilbert and Boulter, 1998) and the difference in learning demand of different scientific concepts (e.g. Adey and Shayer, 1981). These all suggest that there are sequences of conceptual development in science concepts or at least learning gains are better made if, in general, approached in a particular sequence.

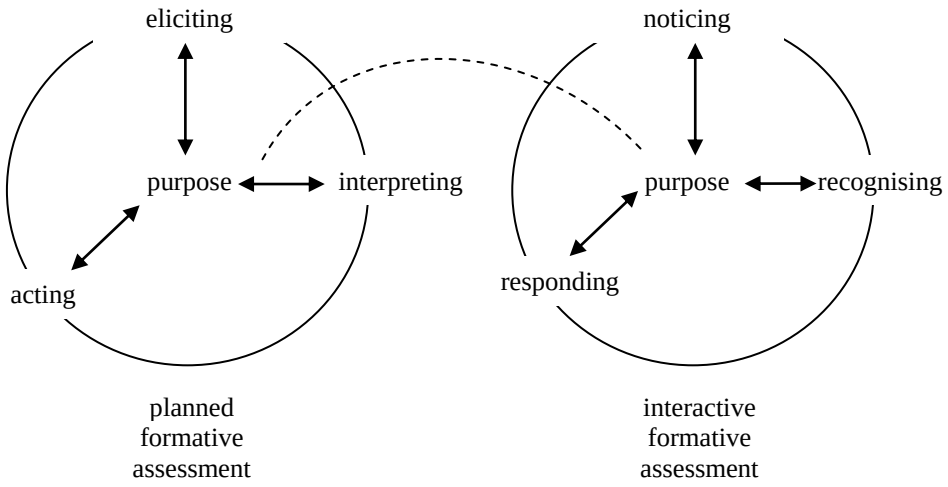
In addition, I would maintain that having a concept of a predetermined sequence of 'general concepts' equips teachers to teach within the Vygotskian concept of ZPD. Just because there is linear progression does not determine that there are behaviourist assumptions, a lot of science education is constructivist through for example, pupils using models and modelling to explain scientific phenomenon (e.g. Coll *et al.*, 2005, Taber, 2006). I assert that a skilled science teacher can work with pupils to judge what they understand and what their next steps (their ZPD) could be within a framework of conceptual development in curriculum science. This appears particularly apparent when

observing teachers who regularly use strategies like the 'Levels Mountain' effectively (Newberry *et al.*, 2005, Grevatt *et al.*, 2007, Newberry and Gilbert, 2007).

Although Torrance and Pryor (1998) do not suggest that teachers should adopt a solely divergent approach (p.154), they do assert that divergent assessment offers a greater scope for positive effects on pupils' learning. It is clear that the LATs do not fit neatly within either the convergent or the divergent assessment features. However, as I have found with my research, the teacher determines the extent to which the LAT is a formative task. The Summative Case is clearly within the description of convergent assessment, but the Formative Case shows more characteristics of divergent assessment.

Bell and Cowie (2001a) also offer a model of two types of formative assessment grounded in the science classroom. They make the distinction of *planned* formative assessment and *interactive* formative assessment. Bell and Cowie constructed a model through their research that was based on the *purpose* of assessment in science. Their descriptions and explanations of these processes are rich, so here I attempt to summarise their main points and then relate these to the use of the LATs. The purpose for them is a balance of social development, personal development and science learning (p.81). In their construction of *planned* formative assessment the purpose of the assessment is central to the relationship between the actions of the teacher (see Figure 6.1): *eliciting* information using specific assessment tasks, *interpreting* the information often using science criterion-referencing, *acting* on that information is the next step that teachers take to enhance the students' learning. This stage is intimately related to the purpose of the assessment and could be focussed three ways of acting: science-referenced, student-referenced or care-referenced.

Figure 6.1 A model of formative assessment (Bell and Cowie, 2001 p. 91)



Their second form of formative assessment is interactive formative assessment, illustrated on the right hand side of Figure 6.1. Here the *purpose* of assessment was to mediate in students' learning with respect to science, social and personal learning. This starts with the teacher *noticing* which they describe as faster than the planned *eliciting* and note that the information gained was ephemeral (verbal, non-verbal and unrecorded). *Recognising* was the process by which teachers through interacting with the students recognised the significance for the pupils' personal, social or science understandings. They link this to Sadler's (1989) argument that qualitative judgements invoke fuzzy criteria which are context driven and the salience of a particular criteria depends on 'what is deemed to be worth noticing' at a particular time (Bell and Cowie, 2001, p.89). Then in the model, comes *responding* the process that happens in response to *noticing* and *recognising* and they equate this to the *acting* aspect of planned formative assessment, but in this case the response is much more immediate.

As with the Torrance and Pryor (1998) model, the LATs do not appear to fit in either of Bell and Cowie's (2001) categories neatly, but are heavily dependent on how the teacher uses the task. On first appearances it is quite clear that the LATs are a planned formative assessment. For example in the Typical and the Transitional cases, it appears

that they are used as a planned assessment task with the whole class, usually there is a delay in time between the assessment and the feedback, the purposes are completely science referenced, responsive to 'getting through the curriculum' and the assessment information was used as an outcome. However, in the Formative Case particularly, but in elements of the Transitional Case, some interactive formative assessment was happening. For example, the teachers were able to respond to pupils' misconceptions quickly and feedback on questions that the pupils had. The teacher in the Transitional Case stated:

9	T:	I just found it interesting because you view
10		it as an assessment, and so you think, oh,
11		I mustn't tell them, but my view on these is
12		that you just explain all lesson. I spend all
13		lesson explaining, and it's very useful for
14		me, because we've been doing particle
15		diagrams the last few lessons, and clearly
16		only a few of them are even accessing
17		that. So that's hugely valuable for me.

Although the LATs do lend themselves to obtaining summative data (a feature of planned formative assessment) they can indeed be *responsive* to pupil learning as opposed to 'getting through the curriculum' and there are several examples of ephemeral assessment information being used as part of the assessment process from my lesson observations.

Bell and Cowie (2001) identify that both forms of assessment rely on teachers' professional knowledge when they are *interpreting* and *acting* or *recognising* and *responding*. The LATs focus only on the development of science concepts and skills, I suggest that teachers need this support and that the other aspects of Bell and Cowie's purpose of assessment of personal and social factors are often more established as part

of their professional knowledge and respond to these as a matter of course. Although I did not study these aspects in my research, I contend the success of the LATs is due to them providing teachers support for the progression in science through levels, something many have less confidence in.

Researchers have also attempted to define other forms of assessment, either sub categories of formative assessment or assessment practices that do not exactly fit within the summative or formative criteria. For example, Carless's (2007) pre-emptive formative assessment:

> seeks particularly to address the problem of timing in feedback processes. I define pre-emptive formative assessment as teacher actions which attempt to clarify student understandings before misconceptions have resulted in ineffective learning or performance and/or loss of marks in assessed tasks. In other words, it is a form of anticipatory feedback in support of student learning. (p. 176)

This is similar to what I have argued above, when comparing the LATs with Bell and Cowie's interactive formative assessment. It occurs to me that the LATs have the potential to be used, as indicated in the Formative and Transitional Cases to pre-empt misconceptions or at least challenge them more or less immediately. Carless (2007) states that:

> pre-emptive formative assessment is about potential for subsequent learning more than current performance (p. 178).

This I observed happening in these cases, where pupils were encouraged to work in their ZPD. Teachers can use the LATs to inform next steps for a class or individuals immediately or the information can be used to inform the next steps of teaching.

This brings me to ipsative assessment, defined by Kelly (2009) as:

> the central criterion of assessment is the individual's own previous levels of attainment. He/she is not assessed against subject attainment targets nor against the performance of peers but against his/her own earlier performance (p. 153)

There are interesting motivational advantages to using assessments in this way, as the pupil is competing against him or herself. Progression up the Level Ladder could easily be framed in this way by teachers trying to encourage their pupils to improve. This type of approach could be seen very much in the case studies where pupils were able to discuss their previous levels that they achieved in LATs. This could be seen in this quote from the Transitional case:

217	Dawn:	I already did it, but I think I was a little bit doubtful
218		whether I did it.
219	Interviewer:	Okay, so you weren't sure whether you'd done
220		it or not.
221	Dawn:	Yeah, because I knew I wrote it, but I thought I
222		did it a little bit wrong, and then the teacher came
223		round and he gave me quite a challenging one.
224	Interviewer:	He gave you a more challenging target?
225	Dawn:	Yeah, and I was like, really pleased with that,
226		because I'm usually quite nervous when I do it, and
227		now I'm not so nervous.
228	Interviewer:	Right.
229	Dawn:	So I know now that the teacher knows that I
230		can do that level target.

This research has led me to consider to what extent the use of a LAT is formative or summative. Unless the lesson is carried out in silent test conditions, the nature of the LATs means that pupils can learn from the task itself, from each other and from their teacher. This would suggest that in the most part, LAT use promotes at least some characteristics of AfL. Harlen (2006) discusses the relationship between formative and summative assessment and asks if there is really a divide or just good assessment (p.

115). To explore this, Harlen (2006) offers a dimension of assessment purposes and practices (Table 6.1).

Table 6.1 A possible dimension of assessment purposes and practices

(adapted from Harlen, 2006, p. 114)

Formative ⟷ Summative				
	Informal formative	Formal formative	Informal summative	Formal summative
Major focus	What are the next steps in learning?		What has been achieved to date?	
Purpose	To inform next steps in learning	To inform next steps in teaching	To monitor progress against plans	To record achievements of individuals
How is evidence collected?	As normal part of class work	Introduced into normal class work	Introduced into normal class work	Separate task or test
Basis of judgment	Student referenced	Student and criterion referenced	Criterion referenced	Criterion referenced
Judged by	Student and teacher	Teacher	Teacher	Teacher or external marker
Action taken	Feedback to students and teacher	Feedback into teaching plans	Feedback into teaching plans	Report to student, parent, other teachers, etc.
Epithet	Assessment for Learning	Matching	Dip stick	Assessment of learning

LATs seem to straddle both domains of assessment, again dependant on teacher's approach. I have shaded the areas in grey where I believe that the LATs map onto these dimensions.

I have attempted to place the LATs within the conceptual frameworks of formative assessment within existing research. The most significant factor in this discussion is that in my research it is not the fact that the LAT is used in the lesson that promotes formative learning experiences, but instead it is the approach of the teacher and the values to assessment held by the pupils. By default the LATs can be formatively convergent and planned formative assessment activities, with all the characteristics described above. However, when a teacher uses the tasks formatively, it does appear some the characteristics of divergent formative assessment and many opportunities for interactive assessment are apparent. There is also scope for the LATs to be used as a form of ipsative assessment.

6.2 Attitudes to the LATs

The first fuzzy generalisation that I offer from my research is that it was the values that teachers and pupils have towards the use of the tasks that influenced the assessment approach that was used. There has been particular interest in making links between issues facing teachers and researchers with AfL and pupils' motivation. Much of this work has stemmed from psychological research into self-theories and achievement goals carried out by Dweck (2000). A particularly relevant study to my research that has emerged from these ideas is that by Alkharusi (2008) who investigated the relationship between students' self-efficacy and achievement goals as a function of teachers' assessment practices in secondary science in Oman. Alkharusi classifies assessment practices into traditional and alternative assessments. Traditional assessments are described as examinations whereas alternative approaches were considered to be such as portfolios, observations, and other performance-based assessments (p.247). The LATs fall into the latter category.

Commonly, achievement goal theory states that there are two types of achievement goals: mastery and performance goals. Mastery goals centre on the development of competence. Students who adopt mastery goals are expected to persist in the face of difficulty, seek challenging tasks, and have high intrinsic motivation. In contrast performance goals centre on the outward showing of competence. In this case, students who adopt performance goals are expected to minimally persist in the face of difficulty, avoid challenging tasks, and have low intrinsic motivation. Alkharusi (2008) draws on a trichotomous framework developed by Middleton and Midgley (1997) that identifies:

(a) mastery goals that focus on improving competence,

(b) performance-approach goals that focus on displaying competence

(c) performance-avoidance goals that focus on avoiding a display of incompetence

Using psychological surveys with pupils to elicit their attitudes, Alkharusi (2008) analysed the pupils' attitudes as individuals and as an aggregate class. His first two findings were that:

> classes with a high emphasis on traditional assessments featuring objective close-ended test items might encourage students to pursue performance-avoidance goals and
>
> classes with a high emphasis on alternative assessments featuring performance-based assessments tend to discourage students from pursuing performance-approach goals.
>
> (p.264).

This aligns with the observation that the LATs in the Summative Case found that pupils were keener to discuss presentation and spent more time on their title (performance-avoidance) than they were pursuing conceptual development or improvement as was more commonly observed in the Formative Case (mastery-goals). Although in my

study, the same type of task was used, it was used differently in each case study, with emphasis on either summative values or formative values.

Alkharusi also found that:

> classroom assessment emphasizing the importance of grades rather than learning and focusing on public rather than private evaluation practices might orient students toward the adoption of performance goals (sic) and
>
> the collective perception of the class members about their classroom assessment environment as being public-oriented tend to orient an individual student within that class toward the adoption of performance-approach goals. (p. 264).

I did not explicitly investigate the differences between private and public assessment approaches. However, the emphasis of grades over learning being related to how pupils value either performance or learning goals is relevant to my study. Finally, Alkharusi concluded that the collective experience of the class members about the classroom assessment climate could be instrumental in having desirable patterns of student achievement motivation. This further supports my claim that is not the fact that LATs are being used by the teacher that means that formative processes are taking place, but it is the way the teacher uses them. This is the assessment culture of the classroom.

6.3 Levels and Levelness

My second fuzzy generalisation was that teachers with a ‘big picture of levels’ may be more likely to use LATs formatively. The Level Ladder guides both teachers and pupils through a framework of progression based on the NCLDs. Attempts by teachers to visualise or generalise the NCLDs were discussed in Chapter 1. Harlen and Winter (2004) explore a range of features of assessment that have the potential to actively support pupils’ learning (p. 393), one of which was using indicators of learning. They highlight the importance of having a framework through which the teacher can think about progression and therefore consider the improvement targets:

> Since the purpose of the assessment is to help development of ideas, skills and attitudes and to use this information to identify next steps in a pupil's learning it is necessary to have the development mapped out, that is, to see the development towards the achievement of goals as a progression. Determining where the pupil is in this progression enables the teacher to identify the 'next step' towards the goals and thus the action needed to help this step to be taken. Although not enough is known to be definitive about the course of development of some of these learning outcomes, it is useful to use what we do know from research and experience to provide a focusing framework. (Harlen and Winter, 2004 p.396)

I have argued that the use of Level Ladders or the Levels Mountain (Newberry *et al.*, 2005) are adaptations of the NCLDs into a useable framework for both teachers and pupils. However, Harlen and Winter (2004) warn that:

> For the purpose of using classroom assessment to help learning, however, there is no need to relate to stages or levels; what is important is to describe how a skill or understanding develops. While teachers and, increasingly, learners have become familiar with the language of levels, it is important to remember that these are a reporting tool, to convey information about progress at set times in a learner's school life, and do not have great value in supporting teachers in making day-to-day decisions about how to move learners on. (p. 403)

This is a valid point and highlights the assumption that the levels are based on theories of conceptual development as discussed in Chapter 1. The alternative approaches that Harlen and Winter suggest is an elaborated list of questions or behaviours observed that indicate a pupil's success at a particular skill.

To be successful at linking learning and assessment as a teacher, first the teacher must have this overall conceptual framework as the teacher demonstrated in the Formative Case with her models of each level and second the pupils themselves must be engaged with these criteria required for each level in an accessible format. I would argue that the LATs do that and that is why they have proved so popular with science teachers. However, I would caveat that claim with the warning that it appears that a majority of teachers using LATs have not recognised this and so end up using the Level Ladder as a mark scheme rather than part of a framework of progression within the subject.

Although I agree with Harlen and Winter's warning that the 'language of levels' is a reporting tool not an everyday assessment tool, I think that teachers who use it as framework can be more effective at formative assessment and that engaging pupils appropriately with this language can be beneficial in AfL and progression. Harlen and Winter do not define their 'language of levels' but this is important in my third generalisation.

My third fuzzy generalisation was that teachers who engage pupils with the notion of 'levelness' may be more likely to improve conceptual development of pupils. This links to Harlen and Winter's research, but is more specifically about pupil engagement with levelness, not just levels. This is perhaps where there is a lack of relevant research.

Effective AfL requires pupils to be aware of, engage with and be involved with the learning objectives and outcomes of the lesson (Black and William 1998a, Bell and Cowie, 2001). According to my analysis and development of the concepts of 'level' and 'levelness' approaches to using the tasks (Table 5.1), teachers are divided in their approach to assessment and this, I argue, has impact on their pupils' values. If in KS3 science there are generic and accessible descriptors that pupils can engage with, they should, in theory be able to have a better understanding of the requirements of their current level and their next level. By having a shared meaning of what it means to be a particular level between pupils and their teacher, it could facilitate progression in science concepts. This is not the 'language of levels' but the 'language of level*ness*.'

Just by using a LAT does not mean the teacher is 'doing AfL'. AfL is a pedagogy made up a particular set of teacher values that are shared with the pupils (Bell and Cowie, 2001a and Harlen and Winter, 2004). Together these create a climate for formative

assessment, features of which were revealed in the Formative Case. Marshall and Drummond (2006) make a distinction between:

> lessons that embody the 'spirit' of AfL and those that conform only to the 'letter'. The nature and sequence of tasks and especially 'high organization based on ideas' appears crucial to the former. This adds a dimension to more familiar formulations of AFL practices. (p. 133)

Although Marshall and Drummond (2006) have a different perspective on what formative assessment should look like in the classroom, there are some similarities in their observations of two types of teacher to my two extremes of pen-portraits (Table 4.2). They too acknowledge that the teacher is the driver of classroom attitude. They describe this as the 'spirit' of AfL:

> It seems... that the beliefs of some teachers map more readily onto what we have called the spirit of AfL. This is partly because they value pupil autonomy and see it as a key goal of their teaching but it also has something to do with how they see the classroom as a site of their own learning. (p. 147).

They elucidate that a key factor here is the teacher's attitude to pupil autonomy. The LATs do lend themselves to pupil autonomy. They are designed so that pupils can tackle them on their own and in their own way within the criteria, but this is dependent on the teacher's approach. However, this 'spirit' of AfL is something more than 'levelness'. There is a different ideology between these two descriptions. Marshall and Drummond (2006) are describing day-to-day activities where the teacher either has a very fixed view about outcomes from that lesson, the 'I am doing and AfL activity' approach or the 'spirit' of AfL where it is pedagogical approach that is embedded in all activities that the teacher does. I contend that the concept of 'levelness' comes mid-way between the two positions. The Marshall and Drummond (2006) view of AfL is very much the ideal approach that it appears that the major commentators want to see happening in schools (e.g. Black and Wiliam, 1998a and 1998b, Harlen and Winter, 2004). In this ideal approach, there would arguably be no need for LAT style activities.

I use the last part of this chapter to discuss, in light of my research and current academic thought, why it is that LATs have been so popular in science lessons.

6.4 To grade or not to grade?

Black and Wiliam (1998a) use an example of a formative approach where students who had just comments and no marks made more progress than pupils who have marks and comments or just marks (Butler, 1988). Also the quality of feedback is bought up as an issue, because it is either managerial and not related to learning (Black and Wiliam, 1998a) or pupils do not understand the feedback or are not provided with support to act on it (Smith and Gorard, 2005a). I argue that even if giving marks is detrimental to achievement, grades and levels are the currency of the classroom culture in England particularly. Politicians, local authorities, head teachers, heads of departments and parents all want to know the grades of each and every pupil (Mansell, 2007). This leads to an emphasis on grades and teachers have tried to make some sense of this and share the requirements or criteria of grades with their pupils. Grades are part of the classroom culture and I suggest that if the teachers and pupils have a notion of 'levelness' then pupils will make progress. This is because then both teachers and pupils are speaking the same language and this is used in meaningful feedback.

The Butler (1988) study was not carried out in a classroom situation and was set up as an experimental investigation. It did yield striking results in that pupils given comments only made significant improvements compared to those given grades alone or surprisingly grades and comments together (which showed a decline in achievement). The main issue here is that giving pupils their grades can be demotivating, particularly for low achievers. Based on this, some schools have decided not to share grades with pupils apart from when reports are given. For example, the school where the Formative Case Study was carried out was trialling this approach. However, the quality of

feedback is an essential factor of making this successful and Smith and Gorard (2005a) question the size of this effect in ecologically valid situations.

Black and Wiliam (1998b) suggested that 'feedback has been shown to improve learning where it gives pupils specific guidance on strengths and weaknesses, preferably without any overall mark' (p. 8). The use of written comments on students' work, rather than marks, was the key method of feedback that differentiated the treatment from the control groups in the present study. Smith and Gorard (2005a) also suggest that it is the *quality* of the comments and feedback given (p.15). As suggested earlier, this should provide students with guidance on how to improve, as well as opportunities and support to understand how to make the improvement.

There remains debate over this aspect of AfL. Smith and Gorard (2005a) claim that in their small-scale study that the strategy of 'no marks' led to the pupils being disadvantaged to the control group (p.4). Black et al (2005) counterclaim that the Smith and Gorard study is flawed because it contains no formative assessment, just no grades being given (p.14). Whether or not the practice of not sharing marks is beneficial in the AfL classroom, it is a difficult feature to do in the current education climate of grades.

Through my research, I observed in all the case studies using 'levels' to communicate success criteria. In some cases, particularly in the Formative Case and partially in the Transitional Case, teachers and pupils were engaged in what I call the 'language of levelness.' The Butler (1988) study indicates that only improvement comments should be given and grades or grades and comments will have a detrimental effect on pupil motivation. However, neither the Butler study nor the Smith and Gorard (2005a) study engaged the pupils with the language of

the grades. When using LATs formatively, both teacher and pupil had a shared understanding of what each level was 'about.' I do accept that the teacher's understanding is more sophisticated than that of the pupils, but having a common language where pupils and teachers can discuss improvements, expectation and progression could improve the quality of the feedback because pupils have understanding of what the levels mean.

There is another issue at this point and that is the exact purpose of the LATs. They were intended to replace summative topic tests and provide a vehicle for more formative assessment in the classroom. They were never intended for everyday use, however, the 'language of levelness' is something I am aware that does happen in some classrooms using the LATs and Levels Mountain. Torrance (2008) investigated advanced level programmes which are all examined through criterion-referencing. His findings raised fears of teachers using 'assessment as learning':

> This might be characterized as a move from assessment of learning, through the currently popular idea of Assessment for Learning, to assessment as learning, where assessment procedures and practices come completely to dominate the learning experience, and 'criteria compliance' comes to replace 'learning'. (p. 284)

This is an interesting issue. The LATs were designed as a way to enhance AfL in the classroom by having explicit criteria referenced to the NCLDs. In doing this, the assessment criteria come to dominate classroom talk. If the NCLDs were founded on good educational theory, then they would perhaps plot an approximate route of progression for science understanding and skills. However, as discussed in Chapter 1, the theoretical foundations of the current NCLDs are somewhat diluted. Therefore, prioritising the 'language of levelness' that I am promoting over the progress of

scientific concepts and skills, could be at the detriment of pupil motivation, success and scientific learning and skills. On the other hand, if the NCLDs are approximated with one or more educational theories in visualisations such as the Levels Mountain, where pupils can regularly refer to notions of levelness to help improve their learning through communication of learning outcomes and improvement targets then surely this satisfies most definitions of AfL. This latter idea would involve pupils being able to recall for example that Level 5 requires you to use an abstract idea (energy, forces, particles or life and cells) to explain simply what they have seen. The research about sharing grades with pupils does raise some negative motivational implications of the practice. In addition, just giving feedback is not in itself sufficient to have positive effects on learning unless the quality of the feedback is of a good enough standard that pupils understand it and can act on it. However, there appears to be no research on the use of levels as a language through which to communicate shared objectives and outcomes, which in turn could led to improved quality of feedback. An implication of such a practice could, if not done carefully result in 'assessment as learning' at the detriment of learning science. With regard to LATs, my current findings would suggest that this implication is not a common approach and that teachers could use the LATs much more formatively which could perhaps improve attainment levels, if not just improved learning in science. However do we want to see this happening in our science classrooms?

6.5 Levels and no-levels culture clash

So it appears that I have highlighted three main attitudes to AfL and the LATs; the 'level'-summative approach, the 'levelness'-formative approach or the 'no-grade'-formative approach. The level-summative attitude is characterised by Table 4.2 pen

portraits, where both teacher and pupil are focussed on outcomes of a task at any given time in relation to a ‘number’, the NCLD.

The ‘levelness’-formative attitude requires both teachers and pupils (to a lesser extent) having an overall view of progression in concepts and skills from the NC. Although the levels are used, the focus is on a meaning associated with the level (e.g. level 5 means that I have to use an abstract concept to explain a phenomenon) and the teachers share the language to make progress in science learning. Although the numbers are shared, they have meaning and therefore can be used to structure learning outcomes, select improvement targets and act on improvement targets.

Finally the ‘no-grade’-formative approach is the ideal of AfL, where teachers and pupils co-construct learning outcomes, pupils have autonomy in their progression and can identify their own and their peers’ learning needs. The teacher acts as a facilitator for learning and improvement, having an overview of the social, science skills and science understanding that he or she can use to challenge pupils in their ZPD. No grades are given and only good quality written or verbal feedback is shared and acted upon by the pupils. This is what was originally championed by Black and Wiliam (1998a) as one way to improve formative feedback and has since been researched and developed by many others (Gardner, 2006).

I suggest that it is at the frustration of the champions of the latter ideology of AfL that the ‘levelness’ approach exists (Harlen, 2006). The approach is at odds with much of the research on grades, pupil motivation and progress (e.g. Butler, 1988). To practising teachers the ‘no-grades-formative’ approach could seem like an ideology that is impossible to meet in the current education climate. However I, like many teachers, wanted to embrace the AfL approach into my teaching and as a result found a half-way

house, the concept of ‘levelness’. Black and William (2003) in their research with classroom teachers recognise the benefits of criteria based on the NCLDs for teacher, peer and self assessment in Mathematics (pp. 50-51). I now explore the possible reasons for the LATs rise in popularity in science education, their future and how teachers may be encouraged towards a more authentic approach to AfL.

6.6 Why are LATs so popular?

The online survey encouraged 106 responses. On the assumption that there about 4000 secondary state schools in England and according to Badger Publishing, over 1000 state schools have purchased the original series of Badger KS3 LATs, it could be assumed that one in four state schools use the LATs. In addition to this, I know that there are online LATs developed by local authorities that a number of schools use. So it is fair to assume at least a quarter of state schools use the level ladder approach to assessing their pupils, maybe as many as a half of schools may use this approach. This means that the sample of the survey is between 5%-10% of schools using the tasks.

The responses to the survey were very positive, with only two teachers asserting that they did not like the tasks. This could of course be an artefact of a self-selecting survey and maybe only those who were passionate about the LATs may have responded. However, it is clear that many teachers use LAT-style activities as alternative forms of assessment and demand for published activities is high. So the LATs are indeed popular and do fulfil a need, though I accept it does not necessarily follow that the activities are a ‘good thing’ for science education *per se*.

Before I started the doctorate I suffered frustration that many science teachers would not engage with levels in the way I felt was necessary for successful science teaching. I remember in 2003 and 2004 running workshops in East Sussex to teach teachers how to

write LATs. It became clear to me that teachers did not want to engage with writing LATs but instead just be given the ones I had written to use in their lessons. The reason I used levels was because it helped me understand a path of progression for teaching KS3 science. I regularly speak to science curriculum managers that teach topics in a more or less random order, often dictated by equipment and resource availability rather than any consideration for conceptual development.

With all this in mind, from the teachers that did respond, I was able to propose a model of how the LATs may be typically used in science lessons by science teachers (Figure 4.1). The most disappointing observation from my perspective was that the LATs were not typically being used in a formative way.

However, a majority of the teachers did perceive the following impact of the tasks on their teaching. First that using the LATs has improved their own understanding of the NCLDs, second that the tasks had improved their use of AfL strategies and third that using the LATs had improved their assessment of their pupils. In addition, the teachers perceived an impact on learning in that a majority of teachers believed that use of the LATs had improved pupils awareness of their current level in science, had improved pupil awareness of how to improve and finally improved pupil learning (Figure 4.1). The scope of this study was not to substantiate these claims quantitatively, but to elucidate the perceptions of teachers who use the LATs. Overall, a majority of teachers who have used the tasks do perceive a shift in their own understanding and use of AfL and believe there has been an impact on their pupils.

The reason that I originally developed the LATs was to help me understand the requirements of each level and share this with my pupils. Being in a performance driven culture at school, where teachers are under pressure to meet targets for each and every

one of the pupils that they teach, I decided to engage with the NCLDs. It made sense to me to share the criteria with my students. When I was teaching I did use the 'language of levelness' with my pupils and I believe that many pupils did develop an understanding of what was expected at the levels relevant to them.

Looking back at this now, I can see that the Level Ladders were a result of competing influences on teaching science at KS3 in the early part of the new millennium. On one hand I was trying to teach pupils science in a way that I felt was engaging, exciting and would prepare them for life and on the other I had to perform on a culture of 'hyperaccountability' (Mansell, 2007). It would be reasonable to argue, that I was somewhat successful as a practitioner at performing in this climate. I became an AST 'specialising' in AfL based on criteria that were linked to these performance driven ideals.

I have always had the ideal in mind that LATs were not required because every science teacher would have an understanding of progression which they could share with their pupils and use tools like the Levels Mountain to visualise the notions of levelness. However, it is clear that the LATs are as much of a scaffold for teachers as they are for pupils.

I am rather critical of teachers not engaging with the notion of levelness, but realistically what motivation is there for them to engage with levelness, when they can succeed through encouraging rote learning for examinations? It is easy to get by on just giving pupils enough information to pass their examinations at their expected levels or grade. Why bother with the bigger picture? These issues have been highlighted by commentators such as Ball (2008), Mansell (2007) and Pring et al. (2009).

The LATs themselves could be representative of the solutions teachers are looking for when they find themselves in an environment of 'hyperaccountability' but also under pressure from the NSS to be 'doing AfL.' In their review of 14-19 education and training, Pring et al. (2009) make a distinction between Assessment for learning (AfL) and Assessment for Accountability (AfA) (p. 81). The LATs may offer a solution that appears to support AfL and satisfy the performance culture demands for numbers, grades and targets. Table 6.1 illustrates how the LATs may be able to straddle both the formative and summative domains. The fact that I developed them in such an environment makes them transferable to teachers who are in a similar position. These factors could well explain the popularity of the LATs.

In the 'levelled-world' of the contemporary science teacher, where teachers need support in understanding the 'big picture' or assessment criteria and are encouraged to do more AfL, it appears that the LATs provide a compromise. Is this the best we can hope for in the current performance driven culture in education? Encouraging teachers and pupils to engage with levels-with-meaning may be the best we can hope for in the current situation. It does raise questions about the use of the LATs. Does the fact that the 'number' of the level exists on the Level Ladder hinder pupils focus on the meaning of the level? Could the Level Ladders be used without numbers and still be useful? As the author of the LATs I do believe that the LATs can have a positive effect on teaching and learning, particularly in engaging pupils and teachers with some notion of progression in science. With this in mind, what is the future of such activities in science education?

6.7 The Future of LATs

So far the discussion of the LATs, my research and the existing research has raised several questions. While doing this, I have had concerns if the LATs are really the best

way to be encouraging AfL practices in science classrooms. In this section, I summarise the purpose of the LATs within science assessment in light of the research discussed above and then consider the future of such assessment tasks.

The overriding theme throughout the analysis of my research and when comparing this to the existing research is that it is not the fact that the tasks are used that makes teaching and learning formative, it is *how* the tasks are used and valued by teachers and pupils. The process of analysis of my findings against the research has helped verify the use of the LATs and also the limitations of such activities. From these analyses it appears that the LATs serve the following purposes:

- A 'formal formative' or 'informal summative' activity
- Used for opportunities for informal formative feedback to classes or individuals
- Used as one AfL strategy amongst many other AfL strategies
- For some teachers, increased understanding of NCLDs
- Improved shared understanding of progression between pupils and teachers
- Provision of improvement targets linked to learning and the National Curriculum

The limitations of the LATs could be:

- If used too often could lead to 'Assessment as Learning'
- Over use of this style of task leading to 'Level Ladder fatigue'
- Teachers using the tasks as purely summative tests may demotivate pupils
- The levels assigned to LATs should not be treated with the same validity, reliability or accuracy of standard summative tests.

In the current political climate of government centralised assessment control, it seems that the LATs were timely and suited a need that pressurized teachers have with juggling both summative performance demands and trying to achieve formative ideals. The LATs are a product of professionals trying to please everyone: line managers, performance management, targets, demands from the LA, demands from the government and the government tests.

These two demands do raise questions for policy. Kelly (2009) describes the tension of summative assessment and ipsative assessment:

> In considering an ipsative style of assessment, as in concentrating our attention on its formative and diagnostic functions, we find ourselves moving away from the notion that it can be achieved by the use of simplistic forms of external, 'pencil-and-paper' tests. If assessment is to have a genuinely valuable educational role, it must be more sophisticated than that. In particular, it must rely more on the assessments made by teachers of their own pupils.' (p.154).

The future of the LATs does lie in the hands of policy makers and government expectations of teachers. A key point that characterises this is that research indicates that sharing grade information with pupils is detrimental to motivation and learning, but the pragmatic issue of dealing with curriculum demands means that teachers feel they need to be generating numbers to assign to pupils. The concept of 'levelness' I have proposed, could be seen as a half-way house to engage with AfL and still meet the demands of the target-driven culture of a school. Without a change in government emphasis, I suspect this will remain.

The recent introduction of Assessing Pupil Progress (APP), developed by the National Strategies and the QCA, has almost been an endorsement for the LAT approach. In fact, work produced by pupils in response to the Badger published LATs has been used (but not credited) as exemplar material in the National Strategies training materials (DCSF, 2009a). The APP assessment criteria are amplifications of the NCLDs and the National Strategies are packaging this as the tool kit for AfL (DCSF, 2009c). They have been developed to move teachers away from teaching science content and emphasise what is known as 'How Science Works' (DCSF, 2009b). The assessment criteria are laid out on an A3 sheet of what are essentially five level ladders. Although currently non-statutory (DCSF, 2009b), there appears to be a significant drive by the National Strategies to

encourage all science departments to be using them. It appears that the LAT approach is here to stay for at least the immediate future.

6.8 LATs and progress in teaching and learning

If I think about my personal development when using the tasks, I started out in the summative zone of collecting pupils work, assigning a level using the criteria and giving the pupils individual improvement targets, rarely providing them with opportunities to actually make the improvements. I then started to build in time for improvement and once I was confident with the NCLDs, I was able to start giving pupils opportunities for self and peer assessment, target setting and giving them time to make those improvements.

To use the LATs effectively, as discussed in the previous sections, I contend that teachers do need to approach the use of the LATs with a 'spirit of AfL' and I argue a notion of 'levelness.' The introduction of the Badger Science KS3 LATs does map out the ideal use of the tasks specifically that teachers should give the opportunity for pupils to make improvements, and in addition it described the notion of levelness and explains how the LATs were developed using generic criteria (Grevatt, 2005a, b and c).

Black and Wiliam (1998b) observe that teachers will not take up evidence based ideas if they are presented as general principles which they have to translate into their everyday practice (p.15). As a result much AfL research has focussed on a component of training teachers to engage with the pedagogic principles (Marshall and Drummond, 2006; Bell and Cowie, 2001; Black et al, 2003). The LATs do lack this component and although I often do INSET sessions for schools or workshops at professional conferences, teachers are not trained routinely in how to use the LATs more formatively. This has led me to consider if the graphical representation (Figure 4.1) could also be a trajectory of teacher

development. Perhaps there needs to be an extension to the x and y axis, where the scaffold of the Level Ladder is withdrawn and the ideal formative assessment approach is integrated into the practice and attitudes of both teachers and pupils. On these extended axes could be a new label, after Marshall and Drummond (2006) 'Formative Spirit'. There are clues to what this might look like, because Scotland and Wales are attempting to structure their curriculum on more formative pedagogies (Daugherty, 2008).

Although the verbal interactions between the teacher and pupils were not recorded *ad verbatim* during the lessons for each case study, I offer some general observations of these interactions. The quality of talk in science lessons has been examined in detail, notably by Lemke (1990) as well as many others. Mercer et al. (2004) make a distinction between teacher-led interactions and peer group interactions (p. 361).

In the formative case study, the quality of teacher-pupil interactions were much better compared to the summative case study. For example, the teacher in the summative case very much left the pupils to their own devices once the LAT was set. There was no scaffolding (Wood et al, 1978) and any talk that did occur was managerial rather than learning focussed. In comparison the quality of interaction in the formative case study was very much focussed on improvement via the application of scientific concepts that were made explicit in the Level Ladder. With respect to peer group interactions, I was able to note down five separate examples during the Formative case lesson, but none in the summative case. The transitional case study was distinct from the Typical case in that the quality of talk between the teacher and pupils was of a better quality. Particularly the focus on conceptual gains rather than mechanistic issues (e.g. title, lay out of task). Mercer et al. (2004) conducted an experimental intervention programme with primary school pupils that promoted 'Thinking Together.' Here it highlights,

amongst other conclusions, that it is the teachers' role to support children in collaborative talk, which can then lead to learning gains (p. 374). This may provide another example of how the LATs may support both teacher-pupil interactions and peer group interactions by providing a scaffold (the Level Ladder) through which they can focus talk on concept development.

What started out as an investigation into the use of LATs by science teachers in England, has raised issues far greater than I had anticipated when I began. The LATs and their popularity with science teachers it appears is a product of the current government pressure on education, the target focussed culture of schools and classrooms and 'hyperaccountability'. The activities that with the best intentions were developed to support teachers develop their 'Assessment for Learning' skills have at worse encouraged teachers to use them for 'Assessment as Learning' and 'Assessment as Accountability' purposes. In the meantime the LATs do offer many opportunities for true formative assessment. It can be convergent, planned or interactive or within a clear framework of progression through science concepts which can support teaching and learning. Perhaps, if the government realises that their pressure on assessment has deskilled teachers, they will allow teachers in England to be trained and improve AfL strategies that concentrate purely on pupils learning science.

Chapter 7 Next Steps and Reflection

The main purpose of this research was to explore the use of LATs, theorise and be used to inform practice. I feel the study has raised questions beyond my original scope, not only lessons for science teachers and educational resource developers, but some insights that may be of interest to policy makers, curriculum designers and assessors of the curriculum.

The questions that have been raised directly from the study itself, starting from the classroom then onto wider issues, include:

- What is best practice for the use of the LATs in science lessons?
- What is the optimum frequency of using these tasks?
- Should levels be shared so explicitly with pupils? Does this practice have any value on learning science?
- How can science teachers be trained to use the LATs more formatively?
- How can teachers encourage pupils to use the LATs for self and peer assessment?
- Why are the NCLDs not based on sound educational theory?
- How will APPs work if there is not general overview of progression? I suspect an similar analysis and critique of the new APP criteria as that carried out by Sainsbury and Simar (1998) of the NCLD would be useful.

Other questions raised by the teachers in the online questionnaire and through the teacher interviews:

- How do I allow enough time to make these effective in class with all the other curriculum pressures?
- What models work best for structuring time for improvements?
- How can we best engage pupils with the Level Ladder?
- How can the tasks be made more accessible to low attaining pupils?

- How does the accuracy and reliability of the LATs compare to topic tests?
- Collaboration between pupils when doing LATs.
- Pupil engagement with targets (Entwistle and Smith, 2002)

As an author of LATs, the research has raised several issues for me. Interestingly throughout this research I have been through times of pride and times of despair regarding the use of the tasks and whether the tasks are actually doing science education any good at all. The main frustration I felt was when I received the online questionnaire feedback and although it was overwhelmingly positive; it was concerning that teachers were in general using the tasks in a summative way. When reading the AfL literature, at times I felt the LATs were missing the purpose of AfL however, I also became aware that there is an AfL ideology and an AfL practical reality in the classroom. The reality admittedly is driven by pressures from government initiatives that are probably misplaced.

Although there were only good intentions of writing the LATs, it has at times warranted me to worry that I had created a monster that was being used in classrooms to the detriment of science education. I have stated that the ideal would be not to have the need for Level Ladders and that teachers would be professionally competent enough to use the NCLDs or some form of progression to guide their pupils' learning. This is not generally the case and it is certainly not apparent from the questionnaire and the majority of the case studies. However, I have taken comfort in the fact that a vast majority of teachers have felt their understanding of levels and AfL has improved since using the LATs. I also took comfort in the extremely competent teachers I observed in the case studies who were using the tasks as they were intended and in ways that I had not even considered!

The key concern for me is developing a good model of professional development for teachers when using the LATs. This is difficult when writing a book as teachers rarely read the introduction that does explain the intended use of the tasks and gives practical information for their use. I am as guilty as the NSS for producing reams of notes that sit in folders on prep room shelves. In the latest 'APP friendly' books I have included a PowerPoint and INSET session guide for curriculum managers to deliver to their teachers. I hope that will help.

Another issue raised was an ethical one concerning anonymity. I referred to Walford's (2005) questioning of whether anonymity of schools and individuals was appropriate. I worked within the university guidelines and kept anonymity for all schools, teachers and pupils who took part in the research. The main concern I had with assigning new names to the pupils was that I felt it was changing their identity and their names meant something to me. When I was deciding on new names for them, I found this difficult as I did not want to infer anything about the pupil with the new name. As it was I decided to go with names in alphabetical order, but it is something I want to think about further.

Methodologically, I feel that my interviewing skills have improved significantly and I enjoyed talking to pupils particularly in this way. However, the methodology of fuzzy educational case studies research has raised the most issues.

I chose to use Bassey's (1999) fuzzy approach because it suited my needs and my concerns about the making of claims from educational research. As I explained in my methodology, I considered this research as a type of formalisation of professional practice development. Since deciding upon and trying this fuzzy approach, the term has cropped up in a range of situations including descriptions of the functions of formative

assessment (e.g. Harlen, 2006). I decided to use Bassey's approach uncritically for my research, but want to take an opportunity to reflect on the process.

Using this method for my research raised some issues for me. The first was deciding on the difference between a fuzzy proposition and a fuzzy generalisation. I was content that I could make statements from my case studies and that I could make tentative claims that if something occurred in the case study, it could indeed happen in general. However, there is always a possibility that the case could have been a 'one off.' Bassey (1999) does not discuss the difference between propositions and generalisations except to say that the latter is more tentative that the former. Although he does illustrate the two types of generalisation in the example case studies, the difference was not made explicit. At one point, it actually felt clumsy trying to discern the difference. I decided in the end that the fuzzy propositions would be the claims that arose from the individual case studies. The fuzzy generalisations came about from further comparative synthesis of the cases and triangulation with the quantitative survey.

The second issue came from Bassey's separation of statistical generalisability and quantitative or fuzzy generalisability. I relished using the mixed method approach, but at times had to consider carefully how I would triangulate using quantitative and qualitative evidence. This was reconcilable. I could not tell if Bassey gave greater precedence to quantitative studies over qualitative studies. I know if I did another ten similar case studies, my theories will evolve and adjust, though I have a fair amount of confidence the general structure would remain the same. Compare this to the quantitative element of the online survey, I can say exactly the same thing, if I had a greater sample, I do not expect my findings to change greatly.

Not unlike the Level Ladders giving pupils and teachers a framework of thinking about levels, Bassey's fuzzy generalisation approach has been my framework within which to think about making claims to knowledge. I feel that in future research, I may be confident enough to have this scaffold removed and have more confidence in my evidence and the claims that I make from it. In future studies I may use similar methods, but also use theoretical frameworks.

With regards to the online questionnaire, it is likely that further analysis of the responses may reveal some correlations. I missed the opportunity of giving the case-study teachers a copy of the questionnaire to complete. This would help me 'place' them within the pen portraits more accurately. In addition I could investigate if there are correlations between the teachers' attitude to the LATs and their existing practice. This may help answer some of the questions I raised in section 4.2.

A disadvantage of working alone on a doctoral study is that there is little opportunity for collaboration. The validity of the coding schemes would have been improved if independently checked. Although I do offer the opportunity for scrutiny of the code, it does not make my claims as valid as they could be.

One further point I would like to make is that the serendipity of finding five cases that were able to be plotted upon my framework should not be underestimated. If I had not found this type of relationship, the thesis and its conclusions may have been very different.

Personal reflection

> "Educational experiences do not leave people as they were.
> People become, in an important sense, different *persons*."
> (Pring, 2004 p. 15)

This thesis is not just a communication of a piece of research, but also represents a significant point in my learning and development as a researcher. When I was a classroom teacher, I read the above quote by Richard Pring and it stuck with me. At the time I related it to my influence as a teacher on the young minds I was trying to educate. However, it also triggered a memory related to the first time I saw stomata on a leaf under a microscope.

I was fourteen years old and it was part of a science lesson that we were given the opportunity to observe the surface of a leaf under a microscope. As a result of seeing these microscopic pores I asked my teacher whether all leaves had these holes, which he confirmed they did. I asked if that included all the leaves on the trees outside and the grass. Again he confirmed this. In that instant my worldview had changed. There was a microscopic world out there and plants were busy doing things with microscopic organelles of which I had previously had no comprehension. By happy coincidence, about a decade later I was studying genetically engineered tobacco plants to unlock some of the secrets of the biochemistry of photosynthesis.

It is perhaps unsurprising that the quote came to mind again when I sat down to write these final few paragraphs. I am a different person intellectually, professionally and personally as a result of embarking on this doctorate in 2004. Professionally, I have moved from being a full time science AST and an author of three books of LATs to become an author of some twenty two books of resources for science teaching who is now a part-time trainee teacher educator within a university. My role has changed and

my professional identity has changed. I have moved from having a very clear identity as a science teacher to a stage I am not exactly sure how to answer when people ask me about my job! I usually settle for part-time teacher trainer and part-time author.

Over the years, the EdD has presented several critical incidences that have shaped who I am as a researcher. The first was that when I started the course, I bought a hardback notebook and wrote on the first page 'Keep it Real'. The second session we had was a philosophical focus about ontology and epistemology. These were to me new words and the first time I had experienced the possibility of relativism. My motto was challenged severely within hours of beginning the course. In addition I was accused of being a 'positivist' by some of my course colleagues and apparently that was not a good thing!

I now relish discussions about ontology and epistemology and have a clearer understanding of my position as a researcher. This has been shaped by realising through reading, lectures and discussions, quite often by realising what ontological and epistemological positions I do not have any desire to explore further as much as those that I do.

The various EdD assignments have without doubt chartered the course for my development as a researcher. From the first time I critically engaged with the Black and Wiliam (1998a) paper to developing my skills as an interviewer to my first small scale evaluation on the use of LATs as part of continued professional development. The feedback from these assignments has contributed to my academic writing and thinking. However, after intermitting for a year, I returned to the course and became friends with five people who were on the course. We set up a support group on *Facebook* where we discussed the course, the issues we were facing, bared our ignorance safely and explored our various professional and academic perspectives. We met to discuss our

progress and have got to know each other personally. After a year of messages accumulating on *Facebook*, during one of our group meals we decided to analyse the thread from out different perspectives. Together we produced a paper and presented in the 'new researchers' at the Society for Research in Higher Education Conference 2008 in Liverpool (Chandler-Grevatt *et al.*, 2008).

Our presentation was well received and gave me an insight into the world of academia. I began to understand how networking, becoming known for a particular area of expertise and relating to your peers is all part of what it is to be an academic. I also now have a heightened awareness of trying to understand what it is to be an academic, compared to when I was developing my career as a teacher where I feel like it just happened to me! I have learnt that I want to take my new career strategically and steadily and on my own terms rather than the rushed acceleration of my teaching career.

Before writing this section, I went away for a long weekend to have a break from the doctorate before I spent the last month completing it. I took a book that was unrelated to my studies, it was called 'Seed to Seed' by Harberd (2007) and is a diary of a plant scientist who had been studying the genetic responses of thale cress to various stimuli in the laboratory for most of his professional life. He then came across a single plant growing in the wild and decided to observe it over one year and explain his observations in a diary through a molecular biological lens. Reading his account made me nostalgic for what appeared to me the relative simplicity of natural science when unpicking DNA and observing molecular responses in plants compared to unpicking the culture of a science classroom.

Over the past twelve years since I first qualified as a science teacher, my ideas have changed about education, teaching and learning; the EdD caused me at times to doubt

everything I believed as a teacher. What has remained the same is my passion for science education, for better science education and for better education of science teachers.

I do not doubt now that the LATs have improved science education to some extent, at least reducing the number of summative topic tests that pupils were doing. It continues to give me pride when I visit schools and see pupils' work on the laboratory walls that was the result of the LATs that I have written. The EdD has been an amazing experience and I feel excited about continuing my education as a researcher.

My next steps are to present aspects of this thesis at conferences, write a paper for publication in an academic journal and become a functional member of the academic community within the University. I would like to extend my research into progression in science learning, teacher development and criteria based assessment. In addition, I would like to investigate the teaching and learning of plant science in schools, something that is recently being sidelined in the curriculum and even within the STEM agenda. Whatever I do, I want to be making a difference to science education at the level of the classroom: changing lives and keeping it real.

References

Adey, P. & Shayer, M. (1981) *Towards a science of science teaching.* London: Heinemann Educational.

Adey, P. & Shayer, M. (1994) *Really raising standards: Cognitive intervention and academic achievement.* London: Routledge.

Alkharusi, H. (2008) Effects of Classroom Assessment Practices on Students. *Educational Assessment,* 13 (4), 243-266.

Anderson, L.W., Krathwohl, D.R., Airasian, P.W. & Samuel, B. (2001) *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Ball, S.J. (2008) *The education debate.* Bristol: The Policy Press.

Bassey, M. (1999) *Case Study Research in Educational Settings. Doing Qualitative Research in Educational Settings.* Buckingham: Taylor & Francis.

Bell, B. & Cowie, B. (2001a) The characteristics of formative assessment in science education. *Science Education,* 85 (5), 536-553.

Bell, B. & Cowie, B. (2001b) *Formative Assessment and Science Education.* Dordrecht: Kluwer Academic Publishers.

Black, P. (unpublished) Pupils' understanding of changes of state.

Black, P. & Wiliam, D. (1998a) Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice,* 5 (1), 7-74.

Black, P. & Wiliam, D. (1998b) Inside the Black Box. *Phi Delta Kappan,* 80 (2), 139-148.

Black, P. & Wiliam, D. (2003) 'In Praise of Educational Research': Formative Assessment, *British Educational Research Journal*, 29(5), 623-637.

Black, P.J., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003) *Assessment for learning: Putting it into practice.* Maidenhead: Open University Press.

Black, P., Harrison, C., Hodgen, J., Marshall, B. & Wiliam, D. (2005) The dissemination of formative assessment: a lesson from, or about, evaluation. *Research Intelligence,* 92, 14-15.

Blackburn, S. (2005) *Truth: A guide.* New York: Oxford University Press.

Bliss, J. and Ogborn. J, (1994), Force and motion from the beginning, *Learning and Instruction,* 4, 7-25.

Bloom, B.S., (1956) *Taxonomy of educational objectives, handbook I: Cognitive domain.* New York: David McKay.

Butler, R. (1988) Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation of interest and performance. *British Journal of Educational Psychology*, 58 (1), 1-14.

Carless, D. (2007) Conceptualizing pre-emptive formative assessment. *Assessment in Education: Principles, Policy & Practice,* 14 (2), 171-184.

Chandler-Grevatt, A., Clayton, S., Creaton, J., Crossland, J., Lefevre, M. & Robertson, S., (2008) Unpicking the Threads: Facebook, Peer Learning and the Professional Doctorate. *Society for Research in Higher Education Postgraduate and Newer Researchers Conference*, Liverpool.

Charmaz, K. (2006) *Constructing grounded theory: A practical guide through qualitative analysis.* London: Sage Publications Ltd.

CHSC, (n.d.) *Thinking Frame* [online]. www.thinkingframe.com [Access Date 19/11/2009].

Clegg, S. (2005) Evidence-based practice in educational research: a critical realist critique of systematic review, *British Journal of Sociology of Education*, 26(3), 415-428.

Cohen, L., Manion, L. & Morrison, K. (2000) *Research methods in education.* 5th ed.: London: RoutledgeFalmer.

Coll, R.K., France, B. & Taylor, I. (2005) The role of models/and analogies in science education: implications from research. *International Journal of Science Education*, 27 (2), 183-198.

Corbin, J.M. & Strauss, A.L. (2008) *Basics of qualitative research.* London: Sage Publications, Inc.

Cox, B. (1995) Cox on the Battle for the English Curriculum. London: Hodder & Stoughton.

Daugherty, R. (2008) Reviewing national curriculum assessment in Wales: how can evidence inform the development of policy? *Cambridge Journal of Education,* 38 (1), 73 - 87.

Davies, P. (2002) Levels of attainment in geography. *Assessment in Education: Principles, Policy and Practice,* 9 (2), 185-204.

DCSF, (2008) *Release: Major Reforms to School Accountability Including an End to Compulsory National Tests for Fourteen Year Olds.* [online]. http://www.dcsf.gov.uk/pns/DisplayPN.cgi?pn_id=2008_0229 [Access Date 23/11/2009].

DCSF, (2009a) *APP: The standards files*. London: Crown Publishing.

DCSF, (2009b) *Assessing Pupil Progress (APP): Science*. London: Crown Publishing.

DCSF, (2009c) *Assessing Pupils' Progress (APP) AfL with APP: developing collaborative school-based approaches.* London: Crown Publishing.

DCSF, (n.d) *Standards Site. Schemes of Work: Science at Key Stage 3.* [online]. http://www.standards.dfes.gov.uk/schemes2/secondary_science/ [Access Date 3/11/2009].

DfEE, (2004) *Science: The National Curriculum for England: Key Stages 1-4.* Revised 2004 edition ed. London: Stationery Office Books.

DfES, (2005a) *Assessment for Learning. Guidance for Senior Leaders.* Key Stage 3 National Strategy, London: Department for Education and Skills.

DfES, (2005b) *Planning and Implementing Progression in Science*. Handout 2.14 "A rough guide to levelness". Key Stage 3 National Strategy, London: Department for Education and Skills.

Donaldson, M.C. (1978) *Children's minds.* London: Taylor & Francis.

Dunne, M., Pryor, J. & Yates, P. (2005) *Becoming a Researcher: A Research Companion for the Social Sciences.* Maidenhead: Open University Press.

Dweck, C.S. (2000) *Self-theories: Their role in motivation, personality, and development.* Hove: Psychology Press.

Entwistle, N. & Smith, C. (2002) Personal understanding and target understanding: Mapping influences on the outcomes of learning. *British Journal of Educational Psychology,* 72 (3), 321-342.

Eraut, M. (1994) *Developing professional knowledge and competence.* London: Routledge.

Franco, A. & Taber, K.S. (2009) Secondary Students' Thinking about Familiar Phenomena: Learners' explanations from a curriculum context where 'particles' is a key idea for organising teaching and learning. *International Journal of Science Education,* 31 (14), 1917-1952.

Gardner, J. (ed.) (2006) *Assessment and Learning,* London: Sage.

Gilbert, J.K. & Boulter, C.J. (1998) Learning science through models and modelling. *International handbook of science education,* 2, 53-66.

Glaser, B.G. & Strauss, A.L. (1977) The discovery of grounded theory: Strategies for qualitative research. New Jersey: Aldine.

Gorard, S. & Taylor, C. (2004) *Combining methods in educational and social research.* Maidenhead: Open University Press.

Grevatt, A. (2005a) *Badger KS3 Science. Level Assessed Tasks for Year 7.* Stevenage: Badger Publishing Limited.

Grevatt, A. (2005b) *Badger KS3 Science. Level Assessed Tasks for Year 8.* Stevenage: Badger Publishing Limited.

Grevatt, A. (2005c) *Badger KS3 Science. Level Assessed Tasks for Year 9.* Stevenage: Badger Publishing Limited.

Grevatt, A. (2008) *Badger Science. Level Assessed Tasks: Year 7 Concepts.* Stevenage: Badger Publishing.

Grevatt, A., Gilbert, J.K. & Newberry, M., (2007) Challenging able science learners through models and modelling. *In* K. Taber ed. *Science Education for Gifted Learners.* Oxon: Routledge, 85–99.

Hammersley, M. (2008) *Questioning qualitative inquiry: critical essays.* London: Sage Publications Ltd.

Harberd, N. (2007) *Seed to Seed: The Secret Life of Plants.* London: Bloomsbury Publishing.

Harlen, W., (2006) On the Relationship Between Assessment for Formative and Summative Purposes. *In* J. Gardner ed. *Assessment and Learning.* London: Sage, 103-117.

Harlen, W. & Winter, J. (2004) The development of assessment for learning: learning from the case of science and mathematics. *Language Testing,* 21 (3), 390.

Hinchliffe, S. & Woodward, K. (2004) *The natural and the social: uncertainty, risk, change.* London: Routledge.

Karmiloff-Smith, A. (1992) *Beyond modularity.* Cambridge, MA: MIT Press.

Kelly, A.V. (2009) *The curriculum: Theory and practice.* London: Sage Publications Ltd.

Kosko, B. (1994) *Fuzzy thinking.* London: Flamingo.

Leach, J. & Scott, P. (2002) Designing and Evaluating Science Teaching Sequences: An Approach Drawing upon the Concept of Learning Demand and a Social Constructivist Perspective on Learning. *Studies in Science Education,* 38, 115-42.

Lemke, J. L. (1990), *Talking Science: Language, Learning and Values.* London: Ablex Publishing Corporation.

Lincoln, Y.S. & Guba, E.G. (1985) *Naturalistic inquiry.* London: Sage.

Mansell, W. (2007) *Education by Numbers: The Tyranny of Testing.* London: Politico's Publishing Ltd.

Marshall, B. & Drummond, J. (2006) How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education,* 21 (2), 133-149.

McNiff, J. (1992) *Action research: Principles and practice.* London: Routledge.

Middleton, M.J. & Midgley, C. (1997) Avoiding the demonstration of lack of ability: An underexplored aspect of goal theory. *Journal of Educational Psychology,* 89, 710-718.

Millar, R. (2003) *Teaching about Energy.* In *Strengthening Teaching and Learning of Energy at Key Stage 3 Science,* additional support pack, Key Stage 3 National Strategy, London: Department for Education and Skills.

Mercer, N. Dawes, L. Wegerif, R. Sams, C. (2004) Reasoning as a scientist: ways of helping children to use language to learn science, *British Educational Research Journal*, 30(3) 359-378.

Newberry, M. & Gilbert, J., (2007) Bringing learners and scientific expertise together. *In* K. Taber ed. *Science Education for Gifted Learners.* London: Routledge.

Newberry, M., Gilbert, J. & Hardcastle, D. (2005) Visualizing progression through the Science curriculum in order to raise standards. *School Science Review,* 86 (316), 87-96.

NSS, (2004) *Assessment for Learning (AfL) subject development materials. In* Dfes ed. 1 ed.: National Secondary Strategy. Key Stage 3 National Strategy, London: Department for Education and Skills.

Pawson, R. & Tilley, N. (1997) *Realistic Evaluation*. London: Sage.

Pring, R. (2004) *Philosophy of Educational Research.* 2nd ed. London: Continuum International Publishing.

Pring, R., Hayward, G., Hodgson, A., Johnson, J., Keep, E., Oancea, A., Rees, G., Spours, K., Wilde, S. & Uk, N.F. (2009) *Education for all: the future of education and training for 14-19 year olds.* Oxon: Routledge.

QCDA, (n.d.) *Assessment: Assessing Pupil Progress* [online]. http://www.qcda.gov.uk/13581.aspx [Access date 10/12/2009]

Robson, C. (2002) *Real world research: A resource for social scientists and practitioner-researchers*. London: Blackwell Publishing.

Sadler, D.R. (1989) Formative assessment and the design of instructional systems. *Instructional science,* 18 (2), 119-144.

Sadler, T.D. (2006) Promoting discourse and argumentation in science teacher education. *Journal of Science Teacher Education,* 17 (4), 323-346.

Sadler, T.D. & Donnelly, L.A. (2006) Socioscientific Argumentation: The effects of content knowledge and morality. *International Journal of Science Education,* 28 (12), 1463 - 1488.

Sainsbury, M. & Sizmur, S. (1998) Level descriptions in the National Curriculum: what kind of criterion referencing is this? *Oxford Review of Education*, 181-193.

Shayer, M. & Adey, P., (2002) Cognitive Acceleration Comes of Age. *In* M. Shayer & P. Adey eds. *Learning intelligence: Cognitive acceleration across the curriculum from 5 to 15 years.* Buckingham: Open University Press.

Smith, E. & Gorard, S. (2005a) 'They don't give us our marks': the role of formative feedback in student progress. *Assessment in Education Principles Policy & Practice,* 12 (1), 21-38.

Smith, E & Gorard, S. (2005b) Putting research into practice: an example from the 'Black Box'. *Research Intelligence,* 91, 4-5.

Spradley, J.P. (1980) *Participant observation.* New York: Holt, Rinehart and Winston.

Stenhouse, L., (1988) Case Study Methods. *In* J.P. Keeves ed. *Educational Research, Methodology, and Measurement: an International Handbook.* Oxford: Pergamon.

Taber, K.S. (2000) Case studies and generalizability: grounded theory and research in science education. *International Journal of Science Education,* 22 (5), 469-487.

Taber, K.S. (2006) Beyond Constructivism: the Progressive Research Programme into Learning Science. *Studies in Science Education,* 42 (1), 125 - 184.

Torgerson, C. & Torgerson, D. (2001) The Need for Randomised Controlled Trials in Educational Research, *British Journal of Educational Studies,* 49(3) 316-328.

Torrance, H. (2008) Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Pedagogy and Practice: Culture and Identities*, 281- 294.

Torrance, H. & Pryor, J. (1998) *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom.* London: Open University Press.

Walford, G. (2005) Research Ethical Guidelines and Anonymity. *International Journal of Research and Method in Education,* 28 (1), 83 - 93.

Wiliam, D., (2001) *Level Best? Levels of attainment in National Curriculum assessment: a report.* London, Association of Teachers and Lecturers.

Wood, D., Wood, H. & Middleton, D. (1978) An Experimental Evaluation of Four Face-to-Face Teaching Strategies, *International Journal of Behavioural Development,* 1(2), 131-147.

Yin, R.K. (1993) *Applications of Case Study Research.* . London and New Delhi: Sage Publications

APPENDIX A

Letter to request observation

Email or letter to the department

Andrew Grevatt
University of Sussex
ajg30@sussex.ac.uk

Dear X

I am writing to you because I understand that you use the *Badger Science Key Stage 3 Level-Assessed Tasks*. I am currently doing a doctorate in education at the University of Sussex and my research is focussed on the evaluation of the use of these tasks.

As part of that research I need to visit science departments that have been using the tasks and gather some information about the use of the *Badger Level-Assessed Tasks*.

I am hoping that you may be able to help me with this research. At your convenience, ideally I would like to spend a day with your department. During the day I would like to observe at least one teacher using the tasks, conduct a 30 minute interview with a teacher who uses the task and possibly interview three pupils who I observe using a task. Also, I would like to collect copies of pupil's work and any data that you could provide on their use.

This visit would not be judgemental and the data gathered would remain confidential. It will be used anonymously as part of my doctoral thesis and in any subsequent publications that result from that. The thesis is due for submission at the end of 2009. The research is being carried out under the supervision of Professor Jo Boaler and in accordance with the University of Sussex Ethical Research Codes.

The information that I gather can be summarised and returned to you in a descriptive form, if you think that would be useful to you.

I appreciate that schools are very hectic places and I will be as flexible as possible to collect the information I need. If you can agree to help me in principle, let me know and we can discuss the details further.

Thank you in anticipation

Yours sincerely

Andy Grevatt

APPENDIX B

Pre-observation letter or email

Andrew Grevatt
University of Sussex
Ajg30@sussex.ac.uk

Dear (Interviewees name)

Re: Badger KS3 Science Level-Assessed Tasks Lesson Observation and Interview.

Thank you for volunteering to be observed and interviewed on Friday 23rd May 2008. This letter explains the purpose and expectations of the lesson observation and post-lesson interview.

Context
I am in my third year of study for a Professional Doctorate in Education (EdD) at the University of Sussex. I am researching the use of criteria based formative assessment in science education. To do this I am carrying out between 10 to 15 case studies of individual teachers of how they use the Badger level-assessed tasks in their lessons. These case studies will take place in a number of schools within the southeast England. From these I hope to select a number of the studies to form the basis of my thesis.

The lesson observation
During the lesson I intend to take extensive written notes focussed on how the task is used by the teacher and the pupils. The observation will be focused on a descriptive account and not to make any judgement on you or your teaching. With your agreement, at appropriate times, I may speak with some of the pupils about their work. Please feel free to introduce me to the pupils as a researcher from the university and let them know that I will be taking notes and may ask them questions. If at any point you wish me to stop observing, you have the right to ask me to leave without giving any reason or any repercussions. During the last 5-10 minutes of the lesson I have a **post-lesson questionnaire** for pupils to complete, the results of which will form part of the case study.

The post lesson interview.
The interview is designed to obtain information about your thoughts about how you used the task, to what extent it supported your teaching and the learning of your pupils. With your permission, I would like to record the interview on a digital Dictaphone. I like to approach these interviews as an 'informal professional chat'. Again, you can withdraw from the interview or refuse to answer any questions at any point without giving me a reason.

The post-lesson pupil interviews
I will be interviewing a group of three pupils after your lesson. The focus on this will be on the work that they produced that lesson. The interview schedule is attached for your information.

Your case study
If you wish, the case study I will build from the all above sources maybe given to you in an unanalysed form about four weeks after my visit. In this form, the study will not be supplied to anyone else without your permission. The case study will be anonymised for use my thesis or related publications. I may want to talk to you again about any interpretations or analysis I make about the case study at a later date, at your convenience.

APPENDIX B

Pre-observation letter or email

Cont…

I appreciate that it can be stressful to be observed and interviewed, but I hope that it will be a useful experience for us both. If you have any further questions or concerns, please do not hesitate to contact me.

Many thanks

Andy Grevatt

cc.
Pupil Interview Schedule
Pupil Post-Lesson Questionnaire

APPENDIX C

Lesson observation page 1

Lesson Observation for doctoral research

School:	LAT:	Date: Time:
Teacher:	Last time LAT used	Class:

Time	T+/- L+/-	Description/Supporting Evidence

APPENDIX C

Lesson observation page 2

Lesson Observation for doctoral research

Teacher: MF Years experience:	Position:	Specialism:

Teacher comments on:
To what extent LAT supported teaching:

To what extent LAT supported learning:

Researcher Lesson Summary and additional notes

List of additional data collected.

- ☐ Class performance data
- ☐ Pupil post-lesson questionnaire
- ☐ Teacher post-lesson interview
- ☐ Pupil post-lesson interview
- ☐ Pupil's work
- ☐ Other

APPENDIX D

Post-lesson Pupil Questionnaire

Level-Assessment Questionnaire

Name:	Age:	Male or Female	Date:	School:

Read each statement. Tick **one** box for each statement.

Statement	Strongly Agree	Agree	Not sure	Disagree	Strongly disagree
I enjoyed doing this task.	☐	☐	☐	☐	☐
I used the Level Ladder to help me to do the task.	☐	☐	☐	☐	☐
I found this task difficult to do.	☐	☐	☐	☐	☐
The Level Ladder showed me what to do to improve.	☐	☐	☐	☐	☐
I did not learn anything from doing this task	☐	☐	☐	☐	☐
The Level Ladder helped me instead of asking the teacher.	☐	☐	☐	☐	☐
I learnt something new while doing this task.	☐	☐	☐	☐	☐

What did you like best about doing this task?	What did you like least about doing this task?

APPENDIX E

Pupil Interview Schedule

Post Lesson Pupil Group Interview

Pre-Interview Notification

- Thank you
- Permission to tape the interview
- Notes may be taken.
- The purpose of the interview – *To find out how Level Assessment Tasks are used and what impact, good or bad, they have on teaching and learning. It is part of a research project that I am doing as a researcher at the University of Sussex.*
- This interview is voluntary they have given their informed consent and that they are free to withdraw at anytime without any repercussions.
- The interview is confidential and you will not be named or recognisable if any parts of the interview are used in my write up.
- Interview will not take no more than 20 minutes.
- No right or wrong answers to any of the questions that I will ask, and that I am interested in are the interviewee's opinions and experiences.
- Feel free to interrupt, ask for clarification or choose not to answer any of the questions.

Audio recording:	Written notes:
W: Name:	Work code: T8/L1/C/PWW
X: Name:	Work code: T8/L1/C/PWX
Y: Name:	Work code: T8/L1/C/PWY
Z: Name	Work code: T8/L1/C/PWZ

	M/F	E	M	S	KS3T	Notes
W						
X						
Y						
Z						

Interview Schedule

Ask pupils to discuss the following statements in relation to their work in front of them.

- What do you think is the best bit about your work?
- Did you learn anything while doing the task, if so what and how?
- What did you find difficult when doing the task?
- What would you do next time to improve your work?

APPENDIX F

Teacher Interview Schedule

Post Lesson Teacher Interview

Pre-Interview Notification

- Thank you
- Permission to tape the interview
- Notes may be taken.
- The purpose of the interview – *To find out how Level Assessment Tasks are used and what impact, good or bad, they have on teaching and learning. It is part of a research project that I am doing as a researcher at the University of Sussex.*
- This interview is voluntary they have given their informed consent and that they are free to withdraw at anytime without any repercussions.
- The interview is confidential and you will not be named or recognisable if any parts of the interview are used in my write up.
- Interview will not take no more than 20 minutes.
- No right or wrong answers to any of the questions that I will ask, and that I am interested in are the interviewee's opinions and experiences.
- Feel free to interrupt, ask for clarification or choose not to answer any of the questions.

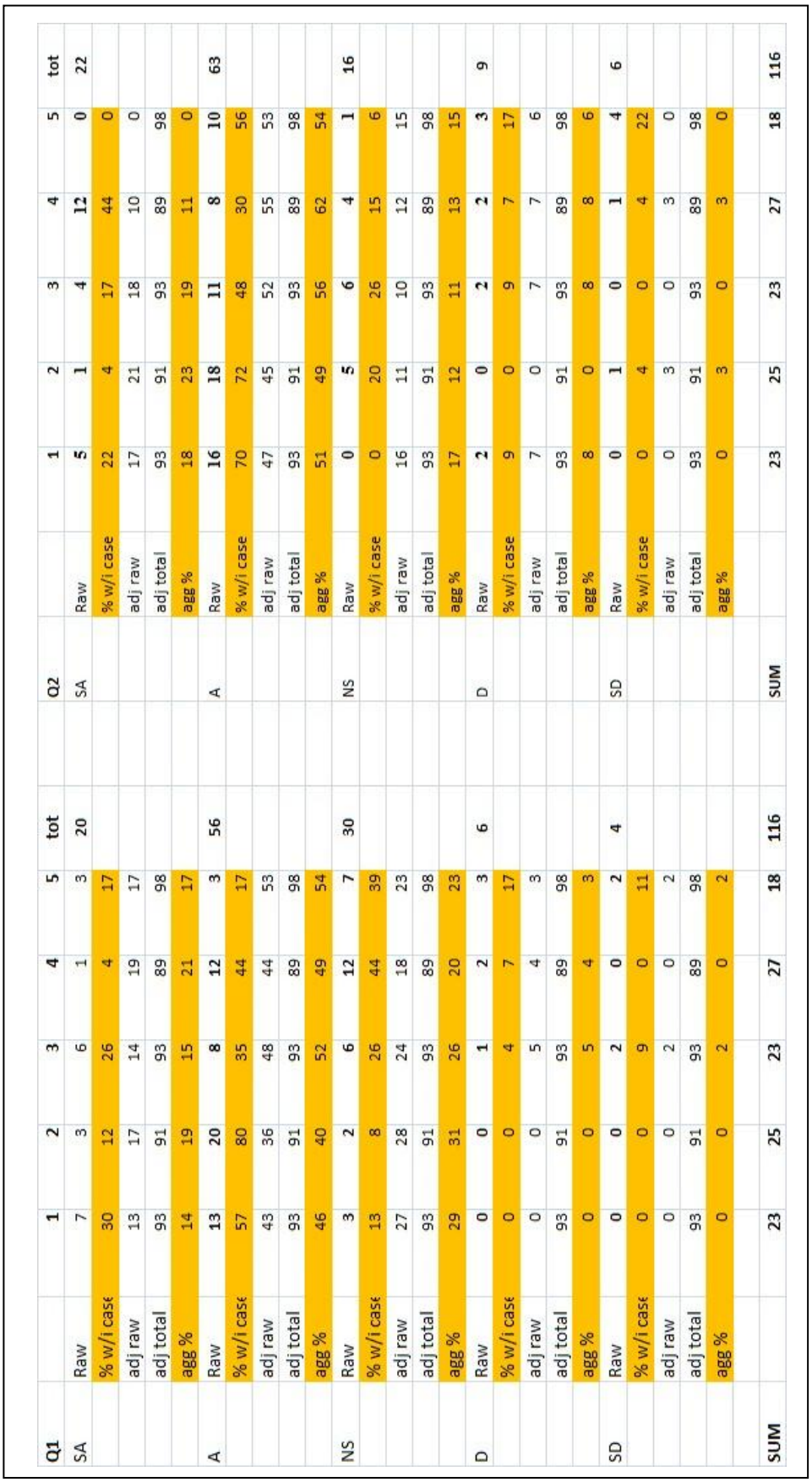
Audio recording:	**Written notes:**
Years Teaching:	Task Used:
No. times used task with this class:	Ability range of group:
How long you have been using the tasks:	
Date: Time:	

Interview Schedule

- Can you give examples of where the Badger task supported your teaching?
- Can you give examples of where the Badger task hindered your teaching?
- Can you give examples of where the Badger Task supported learning?
- Can you give examples of where the Badger Task hindered learning?
- How did you use the Task to plan the lesson?
- How did you adapt the task, why?

APPENDIX G

Adjusted Data Spreadsheet Sample



Q1		1	2	3	4	5	tot	Q2		1	2	3	4	5	tot
SA	Raw	7	3	6	1	3	20	SA	Raw	**5**	**1**	4	**12**	0	22
	% w/i case	30	12	26	4	17			% w/i case	22	4	17	44	0	
	adj raw	13	17	14	19	17			adj raw	17	21	18	10	0	
	adj total	93	91	93	89	98			adj total	93	91	93	89	98	
	agg %	14	19	15	21	17			agg %	18	23	19	11	0	
A	Raw	**13**	20	**8**	**12**	3	56	A	Raw	**16**	**18**	**11**	**8**	**10**	63
	% w/i case	57	80	35	44	17			% w/i case	70	72	48	30	56	
	adj raw	43	36	48	44	53			adj raw	47	45	52	55	53	
	adj total	93	91	93	89	98			adj total	93	91	93	89	98	
	agg %	46	40	52	49	54			agg %	51	49	56	62	54	
NS	Raw	3	2	6	**12**	7	30	NS	Raw	0	**5**	6	4	**1**	16
	% w/i case	13	8	26	44	39			% w/i case	0	20	26	15	6	
	adj raw	27	28	24	18	23			adj raw	16	11	10	12	15	
	adj total	93	91	93	89	98			adj total	93	91	93	89	98	
	agg %	29	31	26	20	23			agg %	17	12	11	13	15	
D	Raw	0	0	**1**	2	3	6	D	Raw	**2**	0	**2**	**2**	**3**	9
	% w/i case	0	0	4	7	17			% w/i case	9	0	9	7	17	
	adj raw	0	0	5	4	3			adj raw	7	0	7	7	6	
	adj total	93	91	93	89	98			adj total	93	91	93	89	98	
	agg %	0	0	5	4	3			agg %	8	0	8	8	6	
SD	Raw	0	0	2	0	2	4	SD	Raw	0	**1**	0	**1**	4	6
	% w/i case	0	0	9	0	11			% w/i case	0	4	0	4	22	
	adj raw	0	0	2	0	2			adj raw	0	3	0	3	0	
	adj total	93	91	93	89	98			adj total	93	91	93	89	98	
	agg %	0	0	2	0	2			agg %	0	3	0	3	0	
SUM		23	25	23	27	18	116	SUM		23	25	23	27	18	116

APPENDIX H

Pupil Questionnaire: Gender Analysis

	Raw	Q1 SA	A	NS	D	SD	Q2 SA	A	NS	D	SD	Q3 SA	A	NS	D	SD
	f	8	28	17	4	2	17	33	4	3	2	1	4	19	27	7
	m	17	26	11	1	0	10	33	1	8	3	1	3	15	32	4
	TOTAL	25	54	28	5	2	27	66	5	11	5	2	7	34	59	11
	%	SA	A	NS	D	SD	SA	A	NS	D	SD	SA	A	NS	D	SD
f	59	14	47	29	7	3	29	56	7	5	3	2	7	32	46	12
m	54	31	48	20	2	0	19	61	2	15	6	2	6	28	59	7
tot	113	22.1	47.8	24.8	4.4	1.8	23.9	58.4	4.4	9.7	4.4	1.8	6.2	30.1	52.2	9.7

Q4 SA	A	NS	D	SD	Q5 SA	A	NS	D	SD	Q6 SA	A	NS	D	SD	Q7 SA	A	NS	D	SD
	10	32	9	4	2	25	27	3	0	3	7	28	12	7	4	12	15	13	6
	9	29	8	5	0	25	17	9	3	1	1	24	22	4	2	1	12	10	21
	19	61	17	9	2	50	44	12	3	4	8	52	34	11	6	13	27	23	24	27
SA	A	NS	D	SD	SA	A	NS	D	SD	SA	A	NS	D	SD	SA	A	NS	D	SD	
	17	54	15	7	3	42	46	5	0	5	12	47	20	12	7	20	25	22	22	10
	17	54	15	9	0	46	31	17	6	2	2	44	41	7	4	2	22	19	20	39
	16.8	54.0	15.0	8.0	1.8	44.2	38.9	10.6	2.7	3.5	7.1	46.0	30.1	9.7	5.3	11.5	23.9	20.4	21.2	23.9

APPENDIX I

Pupil Questionnaire: Gender statistical analysis, Chi Square.

Q1	F	m	Description	
	36	43	Agree	***79***
	17	11	Neutral	***28***
	6	1	Disagree	***7***
	59	***55***	n=	**114**
	f(e)	m(e)		
	40.885965	38.11404	Agree	
	14.491228	13.50877	Neutral	
	3.622807	3.377193	Disagree	
x2=	**0.0691269**	df	2	
Ho	no diff b/w m & f	0.05	5.99	
Hi		0.01	9.21	
		0.001	16.27	

Q2	f	m	Description	
	50	43	Agree	***93***
	4	1	Neutral	***5***
	5	11	Disagree	***16***
	59	***55***	n=	**114**
	f(e)	m(e)		
	48.13157895	44.86842	Agree	
	2.587719298	2.412281	Neutral	
	8.280701754	7.719298	Disagree	
x2=	**0.108500576**	df	2	
Ho	no diff b/w m & f	0.05	5.99	
Hi		0.01	9.21	
		0.001	16.27	

APPENDIX I

Pupil Questionnaire: Gender statistical analysis, Chi Square.

Q3	f		m		Description	
		5		4	Agree	***9***
		19		15	Neutral	***34***
		34		36	Disagree	***70***
		58		***55***	n=	**113**
	f(e)		m(e)			
		4.619469		4.380531	Agree	
		17.45133		16.54867	Neutral	
		35.9292		34.0708	Disagree	
x2=		**0.755938**	df		2	
Ho	no diff b/w m & f			0.05	5.99	
Hi				0.01	9.21	
				0.001	16.27	

Q4	f		m		Description	
		42		38	Agree	***80***
		9		8	Neutral	***17***
		6		5	Disagree	***11***
		57		***51***	n=	**108**
	f(e)		m(e)			
		42.222222		37.77778	Agree	
		8.9722222		8.027778	Neutral	
		5.8055556		5.194444	Disagree	
x2=		**0.9918087**	df		2	
Ho	no diff b/w m & f			0.05	5.99	
Hi				0.01	9.21	
				0.001	16.27	

APPENDIX I

Pupil Questionnaire: Gender statistical analysis, Chi Square.

Q5	f	m	Description	
	52	42	Agree	***94***
	3	9	Neutral	***12***
	3	4	Disagree	***7***
	58	***55***	n=	**113**
	f(e)	m(e)		
	48.24779	45.75221	Agree	
	6.159292	5.840708	Neutral	
	3.59292	3.40708	Disagree	
x2=	**0.126821**	df	2	
Ho	no diff b/w m & f	0.05	5.99	
Hi		0.01	9.21	
		0.001	16.27	

Q6

	f	m	Description	
	35	25	Agree	***60***
	12	22	Neutral	***34***
	11	6	Disagree	***17***
	58	***53***	n=	**111**
	f(e)	m(e)		
	31.351351	28.64865	Agree	
	17.765766	16.23423	Neutral	
	8.8828829	8.117117	Disagree	
x2=	**0.0532609**	df	2	
Ho	no diff b/w m & f	0.05	5.99	
Hi		0.01	9.21	
		0.001	16.27	

APPENDIX I

Pupil Questionnaire: Gender statistical analysis, Chi Square.

Q7

	f		m	Description	
		27	13	Agree	***40***
		13	10	Neutral	***23***
		19	32	Disagree	***51***
		59	***55***	n=	**114**
	f(e)		m(e)		
		20.70175	19.298246	Agree	
		11.90351	11.096491	Neutral	
		26.39474	24.605263	Disagree	
x2=		**0.014443**	df	2	
Ho	no diff b/w m & f		0.05	5.99	
Hi			0.01	9.21	
			0.001	16.27	